

SEQUENCE LISTING

<110> Whitney, Michael  
Sanders, Pamela  
Zeh, Karin

<120> METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF SCREENING ASSAYS

<130> VSD/02-01 PROV US

<160> 64

<170> PatentIn version 3.1

<210> 1

<211> 795

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(795)

<223>  $\beta$ -lactamase

<400> 1

atg agt cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag	48
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	
1 5 10 15	
ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag	96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	
20 25 30	
atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act	144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	
35 40 45	
ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg	192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly	
50 55 60	
caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt	240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val	
65 70 75 80	
gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta	288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val	
85 90 95	
aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc	336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala	
100 105 110	
aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt	384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe	
115 120 125	
ttg cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg	432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro	
130 135 140	
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct	480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro	

145	150	155	160	
gca gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt				528
Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu	165	170	175	
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa				576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys	180	185	190	
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att				624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile	195	200	205	
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca				672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala	210	215	220	
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg				720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr	225	230	235	240
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag				768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu	245	250	255	
ata ggt gcc tca ctg att aag cat tgg				795
Ile Gly Ala Ser Leu Ile Lys His Trp	260	265		
<210>	2			
<211>	265			
<212>	PRT			
<213>	Escherichia coli			
<400>	2			
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln				
1	5	10	15	
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys				
	20	25	30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr				
	35	40	45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly				
	50	55	60	
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val				
65	70	75	80	
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val				
	85	90	95	
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala				
	100	105	110	

Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe  
115 120 125

Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro  
130 135 140

Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro  
145 150 155 160

Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu  
165 170 175

Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys  
180 185 190

Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile  
195 200 205

Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala  
210 215 220

Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr  
225 230 235 240

Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu  
245 250 255

Ile Gly Ala Ser Leu Ile Lys His Trp  
260 265

<210> 3  
<211> 858  
<212> DNA  
<213> Escherichia coli

<220>  
<221> CDS  
<222> (1)..(858)  
<223>  $\beta$ -lactamase

<400> 3  
atg aga att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca 48  
Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala  
1 5 10 15

ttt tgc ctt cct gtt ttt ggt cac cca gaa acg ctg gtg aaa gta aaa 96  
Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys  
20 25 30

gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat 144  
Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp

35	40	45	
ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe 50 55 60			192
cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser 65 70 75 80			240
cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser 85 90 95			288
cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr 100 105 110			336
gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser 115 120 125			384
gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys 130 135 140			432
gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu 145 150 155 160			480
gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg 165 170 175			528
gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu 180 185 190			576
act ggc gaa cta ctt act cta gct tcc cgg caa caa tta ata gac tgg Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp 195 200 205			624
atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro 210 215 220			672
gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser 225 230 235 240			720
cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile 245 250 255			768
gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn 260 265 270			816
aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp 275 280 285			858

<211> 286  
 <212> PRT  
 <213> Escherichia coli

<400> 4

Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala  
 1 5 10 15

Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys  
 20 25 30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp  
 35 40 45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe  
 50 55 60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser  
 65 70 75 80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser  
 85 90 95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr  
 100 105 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser  
 115 120 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys  
 130 135 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu  
 145 150 155 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg  
 165 170 175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu  
 180 185 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp  
 195 200 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro  
 210 215 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser  
 225 230 235 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile  
245 250 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn  
260 265 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp  
275 280 285

<210> 5  
<211> 795  
<212> DNA  
<213> Escherichia coli

<220>  
<221> CDS  
<222> (1)..(795)  
<223>  $\beta$ -lactamase

<400> 5  
atg ggg cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag 48  
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln  
1 5 10 15  
ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag 96  
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys  
20 25 30  
atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act 144  
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr  
35 40 45  
ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gat gac gcc ggg 192  
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly  
50 55 60  
caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt 240  
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val  
65 70 75 80  
gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta 288  
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val  
85 90 95  
aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc 336  
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala  
100 105 110  
aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt 384  
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe  
115 120 125  
ttg cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg 432  
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro  
130 135 140  
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct 480  
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro

145	150	155	160	
gta gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt				528
Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu	165	170	175	
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa				576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys	180	185	190	
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att				624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile	195	200	205	
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca				672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala	210	215	220	
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg				720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr	225	230	235	240
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag				768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu	245	250	255	
ata ggt gcc tca ctg att aag cat tgg				795
Ile Gly Ala Ser Leu Ile Lys His Trp	260	265		
<210>	6			
<211>	265			
<212>	PRT			
<213>	Escherichia coli			
<400>	6			
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	1	5	10	15
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	20	25	30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	35	40	45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly	50	55	60	
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val	65	70	75	80
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val	85	90	95	
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala	100	105	110	

Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe  
115 120 125

Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro  
130 135 140

Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro  
145 150 155 160

Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu  
165 170 175

Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys  
180 185 190

Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile  
195 200 205

Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala  
210 215 220

Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr  
225 230 235 240

Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu  
245 250 255

Ile Gly Ala Ser Leu Ile Lys His Trp  
260 265

<210> 7  
<211> 792  
<212> DNA  
<213> Escherichia coli

<220>  
<221> CDS  
<222> (1)..(792)  
<223> ß-lactamase

<400> 7  
atg gac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg 48  
Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu  
1 5 10 15  
  
ggg gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc 96  
Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile  
20 25 30  
  
ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt 144  
Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe



35	40	45	
aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt att gac gcc ggg caa			192
Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln			
50	55	60	
gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag			240
Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu			
65	70	75	80
tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga			288
Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg			
	85	90	95
gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac			336
Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn			
	100	105	110
tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt ttg			384
Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu			
	115	120	125
cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg gag			432
His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu			
	130	135	140
ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct gta			480
Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val			
	145	150	155
gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act			528
Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr			
	165	170	175
cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt			576
Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val			
	180	185	190
gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct			624
Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala			
	195	200	205
gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca			672
Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala			
	210	215	220
ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg			720
Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr			
	225	230	235
ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag ata			768
Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile			
	245	250	255
ggt gcc tca ctg att aag cat tgg			792
Gly Ala Ser Leu Ile Lys His Trp			
	260		

<210> 8  
 <211> 264  
 <212> PRT  
 <213> Escherichia coli

<400> 8

Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu  
1 5 10 15

Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile  
20 25 30

Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe  
35 40 45

Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln  
50 55 60

Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu  
65 70 75 80

Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg  
85 90 95

Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn  
100 105 110

Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu  
115 120 125

His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu  
130 135 140

Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val  
145 150 155 160

Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr  
165 170 175

Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val  
180 185 190

Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala  
195 200 205

Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala  
210 215 220

Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr  
225 230 235 240

Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile  
245 250 255

Gly Ala Ser Leu Ile Lys His Trp  
260

<210> 9  
<211> 786  
<212> DNA  
<213> Escherichia coli

<220>  
<221> CDS  
<222> (1)..(786)  
<223>  $\beta$ -lactamase

<400> 9  
atg aaa gat gat ttt gca aaa ctt gag gaa caa ttt gat gca aaa ctc 48  
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu  
1 5 10 15  
ggg atc ttt gca ttg gat aca ggt aca aac cgg acg gta gcg tat cgg 96  
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg  
20 25 30  
ccg gat gag cgt ttt gct ttt gct tcg acg att aag gct tta act gta 144  
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val  
35 40 45  
ggc gtg ctt ttg caa cag aaa tca ata gaa gat ctg aac cag aga ata 192  
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile  
50 55 60  
aca tat aca cgt gat gat ctt gta aac tac aac ccg att acg gaa aag 240  
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys  
65 70 75 80  
cac gtt gat acg gga atg acg ctc aaa gag ctt gcg gat gct tcg ctt 288  
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu  
85 90 95  
cga tat agt gac aat gcg gca cag aat ctc att ctt aaa caa att ggc 336  
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly  
100 105 110  
gga cct gaa agt ttg aaa aag gaa ctg agg aag att ggt gat gag gtt 384  
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val  
115 120 125  
aca aat ccc gaa cga ttc gaa cca gag tta aat gaa gtg aat ccg ggt 432  
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly  
130 135 140  
gaa act cag gat acc agt aca gca aga gca ctt gtc aca agc ctt cga 480  
Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg  
145 150 155 160  
gcc ttt gct ctt gaa gat aaa ctt cca agt gaa aaa cgc gag ctt tta 528  
Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu  
165 170 175  
atc gat tgg atg aaa cga aat acc act gga gac gcc tta atc cgt gcc 576  
Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala

180	185	190	
gga gcg gca tca tat gga acc cgg aat gac att gcc atc att tgg ccg			624
Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro			
195	200	205	
cca aaa gga gat cct gtc ggt gtg ccg gac ggt tgg gaa gtg gct gat			672
Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp			
210	215	220	
aaa act gtt ctt gca gta tta tcc agc agg gat aaa aag gac gcc aag			720
Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys			
225	230	235	240
tat gat gat aaa ctt att gca gag gca aca aag gtg gta atg aaa gcc			768
Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala			
245	250	255	
tta aac atg aac ggc aaa			786
Leu Asn Met Asn Gly Lys			
260			
<210> 10			
<211> 262			
<212> PRT			
<213> Escherichia coli			
<400> 10			
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu			
1	5	10	15
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg			
20	25	30	
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val			
35	40	45	
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile			
50	55	60	
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys			
65	70	75	80
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu			
85	90	95	
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly			
100	105	110	
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val			
115	120	125	
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly			
130	135	140	

Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg  
145 150 155 160

Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu  
165 170 175

Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala  
180 185 190

Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro  
195 200 205

Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp  
210 215 220

Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys  
225 230 235 240

Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala  
245 250 255

Leu Asn Met Asn Gly Lys  
260

<210> 11  
<211> 720  
<212> DNA  
<213> Aequorea victoria

<220>  
<221> CDS  
<222> (1)..(720)  
<223> mutant green fluorescent protein

<400> 11  
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg 48  
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu  
1 5 10 15  
gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc 96  
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly  
20 25 30  
gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc 144  
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile  
35 40 45  
tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc 192  
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr  
50 55 60  
ttc tcc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag 240  
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys

65	70	75	80	
cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag				288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu				
	85	90	95	
cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag				336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu				
	100	105	110	
gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc				384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly				
	115	120	125	
atc gac ttc aag gag gac ggc aac atc ctg ggg cac aac ctg gag tac				432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr				
	130	135	140	
aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac				480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn				
	145	150	155	160
ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc				528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser				
	165	170	175	
gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc				576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly				
	180	185	190	
ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg				624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu				
	195	200	205	
agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc				672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe				
	210	215	220	
gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa				720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys				
	225	230	235	
<210> 12				
<211> 239				
<212> PRT				
<213> Aequorea victoria				
<400> 12				
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu				
1 5 10 15				
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly				
20 25 30				
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile				
35 40 45				
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr				
50 55 60				

Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys  
65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu  
85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu  
100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly  
115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr  
130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn  
145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser  
165 170 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly  
180 185 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu  
195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
210 215 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys  
225 230 235

<210> 13  
<211> 690  
<212> DNA  
<213> Anemonia majano

<220>  
<221> CDS  
<222> (1)..(690)  
<223> fluorescent protein

<400> 13  
atg gct ctt tca aac aag ttt atc gga gat gac atg aaa atg acc tac 48  
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr  
1 5 10 15

cat atg gat ggc tgt gtc aat ggg cat tac ttt acc gtc aaa ggt gaa 96  
His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu

20										25										30										
ggc	aac	ggg	aag	cca	tac	gaa	ggg	acg	cag	act	tcg	act	ttt	aaa	gtc	144														
Gly	Asn	Gly	Lys	Pro	Tyr	Glu	Gly	Thr	Gln	Thr	Ser	Thr	Phe	Lys	Val															
35					40					45																				
acc	atg	gcc	aac	ggg	ggg	ccc	ctt	gca	ttc	tcc	ttt	gac	ata	cta	tct	192														
Thr	Met	Ala	Asn	Gly	Gly	Pro	Leu	Ala	Phe	Ser	Phe	Asp	Ile	Leu	Ser															
50					55					60																				
aca	gtg	ttc	aaa	tat	gga	aat	cga	tgc	ttt	act	gcg	tat	cct	acc	agt	240														
Thr	Val	Phe	Lys	Tyr	Gly	Asn	Arg	Cys	Phe	Thr	Ala	Tyr	Pro	Thr	Ser															
65					70					75					80															
atg	ccc	gac	tat	ttc	aaa	caa	gca	ttt	cct	gac	gga	atg	tca	tat	gaa	288														
Met	Pro	Asp	Tyr	Phe	Lys	Gln	Ala	Phe	Pro	Asp	Gly	Met	Ser	Tyr	Glu															
85					90					95																				
agg	act	ttt	acc	tat	gaa	gat	gga	gga	gtt	gct	aca	gcc	agt	tgg	gaa	336														
Arg	Thr	Phe	Thr	Tyr	Glu	Asp	Gly	Gly	Val	Ala	Thr	Ala	Ser	Trp	Glu															
100					105					110																				
ata	agc	ctt	aaa	ggc	aac	tgc	ttt	gag	cac	aaa	tcc	acg	ttt	cat	gga	384														
Ile	Ser	Leu	Lys	Gly	Asn	Cys	Phe	Glu	His	Lys	Ser	Thr	Phe	His	Gly															
115					120					125																				
gtg	aac	ttt	cct	gct	gat	gga	cct	gtg	atg	gcg	aag	aag	aca	act	ggg	432														
Val	Asn	Phe	Pro	Ala	Asp	Gly	Pro	Val	Met	Ala	Lys	Lys	Thr	Thr	Gly															
130					135					140																				
tgg	gac	cca	tct	ttt	gag	aaa	atg	act	gtc	tgc	gat	gga	ata	ttg	aag	480														
Trp	Asp	Pro	Ser	Phe	Glu	Lys	Met	Thr	Val	Cys	Asp	Gly	Ile	Leu	Lys															
145					150					155					160															
ggg	gat	gtc	acc	gcg	ttc	ctc	atg	ctg	caa	gga	ggg	ggc	aat	tac	aga	528														
Gly	Asp	Val	Thr	Ala	Phe	Leu	Met	Leu	Gln	Gly	Gly	Gly	Asn	Tyr	Arg															
165					170					175																				
tgc	caa	ttc	cac	act	tct	tac	aag	aca	aaa	aaa	ccg	gtg	acg	atg	cca	576														
Cys	Gln	Phe	His	Thr	Ser	Tyr	Lys	Thr	Lys	Lys	Pro	Val	Thr	Met	Pro															
180					185					190																				
cca	aac	cat	gtg	gtg	gaa	cat	cgc	att	gcg	agg	acc	gac	ctt	gac	aaa	624														
Pro	Asn	His	Val	Val	Glu	His	Arg	Ile	Ala	Arg	Thr	Asp	Leu	Asp	Lys															
195					200					205																				
ggg	ggc	aac	agt	gtt	cag	ctg	acg	gag	cac	gct	gtt	gca	cat	ata	acc	672														
Gly	Gly	Asn	Ser	Val	Gln	Leu	Thr	Glu	His	Ala	Val	Ala	His	Ile	Thr															
210					215					220																				
tct	gtt	gtc	cct	ttc	tga											690														
Ser	Val	Val	Pro	Phe																										
225																														

<210> 14  
 <211> 229  
 <212> PRT  
 <213> Anemonia majano

<400> 14

Met	Ala	Leu	Ser	Asn	Lys	Phe	Ile	Gly	Asp	Asp	Met	Lys	Met	Thr	Tyr
1				5					10					15	



His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu  
20 25 30

Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val  
35 40 45

Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser  
50 55 60

Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser  
65 70 75 80

Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu  
85 90 95

Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu  
100 105 110

Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly  
115 120 125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly  
130 135 140

Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys  
145 150 155 160

Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg  
165 170 175

Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro  
180 185 190

Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys  
195 200 205

Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr  
210 215 220

Ser Val Val Pro Phe  
225

<210> 15  
<211> 696  
<212> DNA  
<213> Zoanthus sp

<220>

<221> CDS  
 <222> (1)..(696)  
 <223> fluorescent protein

```

<400> 15
atg gct cag tca aag cac ggt cta aca aaa gaa atg aca atg aaa tac      48
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
1                               5                               10                               15

cgt atg gaa ggg tgc gtc gat gga cat aaa ttt gtg atc acg gga gag      96
Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
                20                25                30

ggc att gga tat ccg ttc aaa ggg aaa cag gct att aat ctg tgt gtg      144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
                35                40                45

gtc gaa ggt gga cca ttg cca ttt gcc gaa gac ata ttg tca gct gcc      192
Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
                50                55                60

ttt aac tac gga aac agg gtt ttc act gaa tat cct caa gac ata gtt      240
Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
65                70                75                80

gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg gac agg tct      288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser
                85                90                95

ttt ctc ttt gag gat gga gca gtt tgc ata tgt aat gca gat ata aca      336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
                100                105                110

gtg agt gtt gaa gaa aac tgc atg tat cat gag tcc aaa ttt tat gga      384
Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
                115                120                125

gtg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca gat aac      432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
                130                135                140

tgg gag cca tcc tgc gag aag atc ata cca gta cct aag cag ggg ata      480
Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
145                150                155                160

ttg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt      528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165                170                175

tta cgg tgc caa ttc gac aca gtt tac aaa gca aag tct gtg cca aga      576
Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
                180                185                190

aag atg ccg gac tgg cac ttc atc cag cat aag ctc acc cgt gaa gac      624
Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
                195                200                205

cgc agc gat gct aag aat cag aaa tgg cat ctg aca gaa cat gct att      672
Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
                210                215                220

gca tcc gga tct gca ttg ccc tga      696
Ala Ser Gly Ser Ala Leu Pro
  
```

225

230

<210> 16  
 <211> 231  
 <212> PRT  
 <213> Zoanthus sp

<400> 16

Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr  
 1 5 10 15

Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu  
 20 25 30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val  
 35 40 45

Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala  
 50 55 60

Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val  
 65 70 75 80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser  
 85 90 95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr  
 100 105 110

Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly  
 115 120 125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn  
 130 135 140

Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile  
 145 150 155 160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg  
 165 170 175

Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg  
 180 185 190

Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp  
 195 200 205

Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile  
 210 215 220

Ala Ser Gly Ser Ala Leu Pro  
225 230

<210> 17  
<211> 696  
<212> DNA  
<213> Zoanthus sp

<220>  
<221> CDS  
<222> (1)..(696)  
<223> fluorescent protein

<400> 17  
atg gct cat tca aag cac ggt cta aaa gaa gaa atg aca atg aaa tac 48  
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr  
1 5 10 15  
  
cac atg gaa ggg tgc gtc aac gga cat aaa ttt gtg atc acg ggc gaa 96  
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu  
20 25 30  
  
ggc att gga tat ccg ttc aaa ggg aaa cag act att aat ctg tgt gtg 144  
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val  
35 40 45  
  
atc gaa ggg gga cca ttg cca ttt tcc gaa gac ata ttg tca gct ggc 192  
Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly  
50 55 60  
  
ttt aag tac gga gac agg att ttc act gaa tat cct caa gac ata gta 240  
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val  
65 70 75 80  
  
gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg ggc agg tct 288  
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser  
85 90 95  
  
ttt ctc ttt gag gat gga gca gtc tgc ata tgc aat gta gat ata aca 336  
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr  
100 105 110  
  
gtg agt gtc aaa gaa aac tgc att tat cat aag agc ata ttt aat gga 384  
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly  
115 120 125  
  
atg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca act aac 432  
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn  
130 135 140  
  
tgg gaa gca tcc tgc gag aag atc atg cca gta cct aag cag ggg ata 480  
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile  
145 150 155 160  
  
ctg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt 528  
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Lys Asp Gly Gly Arg  
165 170 175  
  
tac cgg tgc cag ttc gac aca gtt tac aaa gca aag tct gtg cca agt 576  
Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser

180	185	190	
aag atg ccg gag tgg cac ttc atc cag cat aag ctc ctc cgt gaa gac			624
Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp			
195	200	205	
cgc agc gat gct aag aat cag aag tgg cag ctg aca gag cat gct att			672
Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile			
210	215	220	
gca ttc cct tct gcc ttg gcc tga			696
Ala Phe Pro Ser Ala Leu Ala			
225	230		
<210> 18			
<211> 231			
<212> PRT			
<213> Zoanthus sp			
<400> 18			
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr			
1	5	10	15
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu			
20	25	30	
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val			
35	40	45	
Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly			
50	55	60	
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val			
65	70	75	80
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser			
85	90	95	
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr			
100	105	110	
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly			
115	120	125	
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn			
130	135	140	
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile			
145	150	155	160
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg			
165	170	175	

Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser  
180 185 190

Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp  
195 200 205

Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile  
210 215 220

Ala Phe Pro Ser Ala Leu Ala  
225 230

<210> 19  
<211> 699  
<212> DNA  
<213> Discosoma striata

<220>  
<221> CDS  
<222> (1)..(699)  
<223> fluorescent protein

<400> 19  
atg agt tgt tcc aag agt gtg atc aag gaa gaa atg ttg atc gat ctt 48  
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu  
1 5 10 15  
  
cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa 96  
His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys  
20 25 30  
  
gga aaa gga cag cct aat gaa ggc acc aat acc gtc acg ctc gag gtt 144  
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val  
35 40 45  
  
acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa 192  
Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln  
50 55 60  
  
ttt cag tat gga aac aag gca ttt gtc cac cac cct gac aac ata cat 240  
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His  
65 70 75 80  
  
gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cgg tcc 288  
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser  
85 90 95  
  
atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt 336  
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser  
100 105 110  
  
ttg aca ggc aac tgt ttc tac tac gac atc aag ttc act ggc ttg aac 384  
Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn  
115 120 125  
  
ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa 432  
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu

130	135	140	
ccg agc act gag cgt ttg tat cct cgt gat ggt gtg ttg ata gga gac			480
Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp			
145	150	155	160
atc cat cat gct ctg aca gtt gaa gga ggt ggt cat tac gca tgt gac			528
Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp			
	165	170	175
att aaa act gtt tac agg gcc aag aag gcc gcc ttg aag atg cca ggg			576
Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly			
	180	185	190
tat cac tat gtt gac acc aaa ctg gtt ata tgg aac aac gac aaa gaa			624
Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu			
	195	200	205
ttc atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cac cat ccg			672
Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro			
	210	215	220
ttc tat gag cca aag aag gat aag taa			699
Phe Tyr Glu Pro Lys Lys Asp Lys			
225	230		
<210>	20		
<211>	232		
<212>	PRT		
<213>	Discosoma striata		
<400>	20		
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu			
1	5	10	15
His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys			
	20	25	30
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val			
	35	40	45
Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln			
	50	55	60
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His			
65	70	75	80
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser			
	85	90	95
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser			
	100	105	110
Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn			
	115	120	125

Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu  
130 135 140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp  
145 150 155 160

Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp  
165 170 175

Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly  
180 185 190

Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu  
195 200 205

Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro  
210 215 220

Phe Tyr Glu Pro Lys Lys Asp Lys  
225 230

<210> 21  
<211> 678  
<212> DNA  
<213> Discosoma sp "red"

<220>  
<221> CDS  
<222> (1)..(678)  
<223> fluorescent protein

<400> 21  
atg agg tct tcc aag aat gtt atc aag gag ttc atg agg ttt aag gtt 48  
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val  
1 5 10 15

cgc atg gaa gga acg gtc aat ggg cac gag ttt gaa ata gaa ggc gaa 96  
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu  
20 25 30

gga gag ggg agg cca tac gaa ggc cac aat acc gta aag ctt aag gta 144  
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val  
35 40 45

acc aag ggg gga cct ttg cca ttt gct tgg gat att ttg tca cca caa 192  
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln  
50 55 60

ttt cag tat gga agc aag gta tat gtc aag cac cct gcc gac ata cca 240  
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro  
65 70 75 80

gac tat aaa aag ctg tca ttt cct gaa gga ttt aaa tgg gaa agg gtc 288  
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val



85										90					95					
atg	aac	ttt	gaa	gac	ggg	ggc	gtc	gtt	act	gta	acc	cag	gat	tcc	agt	336				
Met	Asn	Phe	Glu	Asp	Gly	Gly	Val	Val	Thr	Val	Thr	Gln	Asp	Ser	Ser					
			100						105			110								
ttg	cag	gat	ggc	tgt	ttc	atc	tac	aag	gtc	aag	ttc	att	ggc	gtg	aac	384				
Leu	Gln	Asp	Gly	Cys	Phe	Ile	Tyr	Lys	Val	Lys	Phe	Ile	Gly	Val	Asn					
			115						120			125								
ttt	cct	tcc	gat	gga	cct	gtt	atg	caa	aag	aag	aca	atg	ggc	tgg	gaa	432				
Phe	Pro	Ser	Asp	Gly	Pro	Val	Met	Gln	Lys	Lys	Thr	Met	Gly	Trp	Glu					
			130						135			140								
gcc	agc	act	gag	cgt	ttg	tat	cct	cgt	gat	ggc	gtg	ttg	aaa	gga	gag	480				
Ala	Ser	Thr	Glu	Arg	Leu	Tyr	Pro	Arg	Asp	Gly	Val	Leu	Lys	Gly	Glu					
145					150						155			160						
att	cat	aag	gct	ctg	aag	ctg	aaa	gac	ggg	ggg	cat	tac	cta	gtt	gaa	528				
Ile	His	Lys	Ala	Leu	Lys	Leu	Lys	Asp	Gly	Gly	His	Tyr	Leu	Val	Glu					
			165						170			175								
ttc	aaa	agt	att	tac	atg	gca	aag	aag	cct	gtg	cag	cta	cca	ggg	tac	576				
Phe	Lys	Ser	Ile	Tyr	Met	Ala	Lys	Lys	Pro	Val	Gln	Leu	Pro	Gly	Tyr					
			180						185			190								
tac	tat	gtt	gac	tcc	aaa	ctg	gat	ata	aca	agc	cac	aac	gaa	gac	tat	624				
Tyr	Tyr	Val	Asp	Ser	Lys	Leu	Asp	Ile	Thr	Ser	His	Asn	Glu	Asp	Tyr					
			195						200			205								
aca	atc	gtt	gag	cag	tat	gaa	aga	acc	gag	gga	cgc	cac	cat	ctg	ttc	672				
Thr	Ile	Val	Glu	Gln	Tyr	Glu	Arg	Thr	Glu	Gly	Arg	His	His	Leu	Phe					
			210						215			220								
ctt	taa														678					
Leu																				
225																				
<210> 22																				
<211> 225																				
<212> PRT																				
<213> Discosoma sp "red"																				
<400> 22																				
Met	Arg	Ser	Ser	Lys	Asn	Val	Ile	Lys	Glu	Phe	Met	Arg	Phe	Lys	Val					
1				5					10			15								
Arg	Met	Glu	Gly	Thr	Val	Asn	Gly	His	Glu	Phe	Glu	Ile	Glu	Gly	Glu					
			20						25			30								
Gly	Glu	Gly	Arg	Pro	Tyr	Glu	Gly	His	Asn	Thr	Val	Lys	Leu	Lys	Val					
			35						40			45								
Thr	Lys	Gly	Gly	Pro	Leu	Pro	Phe	Ala	Trp	Asp	Ile	Leu	Ser	Pro	Gln					
			50						55			60								
Phe	Gln	Tyr	Gly	Ser	Lys	Val	Tyr	Val	Lys	His	Pro	Ala	Asp	Ile	Pro					
65					70						75			80						

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val  
85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser  
100 105 110

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn  
115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu  
130 135 140

Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu  
145 150 155 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu  
165 170 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr  
180 185 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr  
195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe  
210 215 220

Leu  
225

<210> 23  
<211> 801  
<212> DNA  
<213> Clavularia sp

<220>  
<221> CDS  
<222> (1)..(801)  
<223> fluorescent protein

<400> 23  
atg aag tgt aaa ttt gtg ttc tgc ctg tcc ttc ttg gtc ctc gcc atc 48  
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile  
1 5 10 15

aca aac gcg aac att ttt ttg aga aac gag gct gac tta gaa gag aag 96  
Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys  
20 25 30

aca ttg aga ata cca aaa gct cta acc acc atg ggt gtg att aaa cca 144  
Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro

35	40	45	
gac atg aag att aag ctg aag atg gaa gga aat gta aac ggg cat gct Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala 50 55 60			192
ttt gtg atc gaa gga gaa gga gaa gga aag cct tac gat ggg aca cac Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His 65 70 75 80			240
act tta aac ctg gaa gtg aag gaa ggt gcg cct ctg cct ttt tct tac Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr 85 90 95			288
gat atc ttg tca aac gcg ttc cag tac gga aac aga gca ttg aca aaa Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys 100 105 110			336
tac cca gac gat ata gca gac tat ttc aag cag tcg ttt ccc gag gga Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly 115 120 125			384
tat tcc tgg gaa aga acc atg act ttt gaa gac aaa ggc att gtc aaa Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys 130 135 140			432
gtg aaa agt gac ata agc atg gag gaa gac tcc ttt atc tat gaa att Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile 145 150 155 160			480
cgt ttt gat ggg atg aac ttt cct ccc aat ggt ccg gtt atg cag aaa Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys 165 170 175			528
aaa act ttg aag tgg gaa cca tcc act gag att atg tac gtg cgt gat Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp 180 185 190			576
gga gtg ctg gtc gga gat att agc cat tct ctg ttg ctg gag gga ggt Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly 195 200 205			624
ggc cat tac cga tgt gac ttc aaa agt att tac aaa gca aaa aaa gtt Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val 210 215 220			672
gtc aaa ttg cca gac tat cac ttt gtg gac cat cgc att gag atc ttg Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu 225 230 235 240			720
aac cat gac aag gat tac aac aaa gta acg ctg tat gag aat gca gtt Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val 245 250 255			768
gct cgc tat tct ttg ctg cca agt cag gcc tag Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala 260 265			801
<210> 24			
<211> 266			
<212> PRT			
<213> Clavularia sp			

<400> 24

Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile  
1 5 10 15

Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys  
20 25 30

Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro  
35 40 45

Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala  
50 55 60

Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His  
65 70 75 80

Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr  
85 90 95

Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys  
100 105 110

Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly  
115 120 125

Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys  
130 135 140

Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile  
145 150 155 160

Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys  
165 170 175

Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp  
180 185 190

Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly  
195 200 205

Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val  
210 215 220

Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu  
225 230 235 240

Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val  
245 250 255

Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala  
260 265

<210> 25  
<211> 881  
<212> PRT  
<213> *Saccharomyces cerevisiae*

<400> 25

Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu  
1 5 10 15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu  
20 25 30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro  
35 40 45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu  
50 55 60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile  
65 70 75 80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu  
85 90 95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala  
100 105 110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser  
115 120 125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu  
130 135 140

Thr Val Ser Ile Asp Ser Ala Ala His His Asp Asn Ser Thr Ile Pro  
145 150 155 160

Leu Asp Phe Met Pro Arg Asp Ala Leu His Gly Phe Asp Trp Ser Glu  
165 170 175

Glu Asp Asp Met Ser Asp Gly Leu Pro Phe Leu Lys Thr Asp Pro Asn  
180 185 190

Asn Asn Gly Phe Phe Gly Asp Gly Ser Leu Leu Cys Ile Leu Arg Ser  
195 200 205

Ile Gly Phe Lys Pro Glu Asn Tyr Thr Asn Ser Asn Val Asn Arg Leu  
210 215 220

Pro Thr Met Ile Thr Asp Arg Tyr Thr Leu Ala Ser Arg Ser Thr Thr  
225 230 235 240

Ser Arg Leu Leu Gln Ser Tyr Leu Asn Asn Phe His Pro Tyr Cys Pro  
245 250 255

Ile Val His Ser Pro Thr Leu Met Met Leu Tyr Asn Asn Gln Ile Glu  
260 265 270

Ile Ala Ser Lys Asp Gln Trp Gln Ile Leu Phe Asn Cys Ile Leu Ala  
275 280 285

Ile Gly Ala Trp Cys Ile Glu Gly Glu Ser Thr Asp Ile Asp Val Phe  
290 295 300

Tyr Tyr Gln Asn Ala Lys Ser His Leu Thr Ser Lys Val Phe Glu Ser  
305 310 315 320

Gly Ser Ile Ile Leu Val Thr Ala Leu His Leu Leu Ser Arg Tyr Thr  
325 330 335

Gln Trp Arg Gln Lys Thr Asn Thr Ser Tyr Asn Phe His Ser Phe Ser  
340 345 350

Ile Arg Met Ala Ile Ser Leu Gly Leu Asn Arg Asp Leu Pro Ser Ser  
355 360 365

Phe Ser Asp Ser Ser Ile Leu Glu Gln Arg Arg Arg Ile Trp Trp Ser  
370 375 380

Val Tyr Ser Trp Glu Ile Gln Leu Ser Leu Leu Tyr Gly Arg Ser Ile  
385 390 395 400

Gln Leu Ser Gln Asn Thr Ile Ser Phe Pro Ser Ser Val Asp Asp Val  
405 410 415

Gln Arg Thr Thr Thr Gly Pro Thr Ile Tyr His Gly Ile Ile Glu Thr  
420 425 430

Ala Arg Leu Leu Gln Val Phe Thr Lys Ile Tyr Glu Leu Asp Lys Thr  
435 440 445

Val Thr Ala Glu Lys Ser Pro Ile Cys Ala Lys Lys Cys Leu Met Ile  
450 455 460

Cys Asn Glu Ile Glu Glu Val Ser Arg Gln Ala Pro Lys Phe Leu Gln  
465 470 475 480

Met Asp Ile Ser Thr Thr Ala Leu Thr Asn Leu Leu Lys Glu His Pro  
485 490 495

Trp Leu Ser Phe Thr Arg Phe Glu Leu Lys Trp Lys Gln Leu Ser Leu  
500 505 510

Ile Ile Tyr Val Leu Arg Asp Phe Phe Thr Asn Phe Thr Gln Lys Lys  
515 520 525

Ser Gln Leu Glu Gln Asp Gln Asn Asp His Gln Ser Tyr Glu Val Lys  
530 535 540

Arg Cys Ser Ile Met Leu Ser Asp Ala Ala Gln Arg Thr Val Met Ser  
545 550 555 560

Val Ser Ser Tyr Met Asp Asn His Asn Val Thr Pro Tyr Phe Ala Trp  
565 570 575

Asn Cys Ser Tyr Tyr Leu Phe Asn Ala Val Leu Val Pro Ile Lys Thr  
580 585 590

Leu Leu Ser Asn Ser Lys Ser Asn Ala Glu Asn Asn Glu Thr Ala Gln  
595 600 605

Leu Leu Gln Gln Ile Asn Thr Val Leu Met Leu Leu Lys Lys Leu Ala  
610 615 620

Thr Phe Lys Ile Gln Thr Cys Glu Lys Tyr Ile Gln Val Leu Glu Glu  
625 630 635 640

Val Cys Ala Pro Phe Leu Leu Ser Gln Cys Ala Ile Pro Leu Pro His  
645 650 655

Ile Ser Tyr Asn Asn Ser Asn Gly Ser Ala Ile Lys Asn Ile Val Gly  
660 665 670

Ser Ala Thr Ile Ala Gln Tyr Pro Thr Leu Pro Glu Glu Asn Val Asn  
675 680 685

Asn Ile Ser Val Lys Tyr Val Ser Pro Gly Ser Val Gly Pro Ser Pro  
690 695 700

Val Pro Leu Lys Ser Gly Ala Ser Phe Ser Asp Leu Val Lys Leu Leu  
705 710 715 720

Ser Asn Arg Pro Pro Ser Arg Asn Ser Pro Val Thr Ile Pro Arg Ser  
725 730 735

Thr Pro Ser His Arg Ser Val Thr Pro Phe Leu Gly Gln Gln Gln Gln  
740 745 750

Leu Gln Ser Leu Val Pro Leu Thr Pro Ser Ala Leu Phe Gly Gly Ala  
755 760 765

Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe Thr  
770 775 780

Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr Asn  
785 790 795 800

Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp Asn  
805 810 815

Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn Asn  
820 825 830

Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn Ala  
835 840 845

Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp Val  
850 855 860

Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys Lys  
865 870 875 880

Glu

<210> 26

<211> 202

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 26

Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg  
1 5 10 15

Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala  
20 25 30

Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys  
35 40 45



Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg  
50 55 60

Gly Ile Arg Leu Leu Gln Glu Glu Glu Glu Gly Leu Pro Leu Val Gly  
65 70 75 80

Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly  
85 90 95

His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu  
100 105 110

Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly  
115 120 125

Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val  
130 135 140

Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys  
145 150 155 160

Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro  
165 170 175

Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala  
180 185 190

Val Gly Val Ile Arg Asn Gly Asp Trp Leu  
195 200

<210> 27  
<211> 17  
<212> DNA  
<213> unidentified

<220>  
<221> misc\_feature  
<222> (4)..(4)  
<223> any purine

<220>  
<221> misc\_feature  
<222> (1)..(17)  
<223> DNA sequence that is bound by GAL4

<220>  
<221> misc\_feature  
<222> (5)..(5)  
<223> any nucleotide

<220>

<221> misc\_feature  
<222> (6)..(6)  
<223> any nucleotide

<220>  
<221> misc\_feature  
<222> (7)..(7)  
<223> any purine

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> any pyrimidine

<220>  
<221> misc\_feature  
<222> (11)..(11)  
<223> any pyrimidine

<220>  
<221> misc\_feature  
<222> (10)..(10)  
<223> any nucleotide

<220>  
<221> misc\_feature  
<222> (12)..(12)  
<223> any nucleotide

<220>  
<221> misc\_feature  
<222> (14)..(14)  
<223> any nucleotide

<400> 27  
cggnnnncnn nncnccg

17

<210> 28  
<211> 14  
<212> DNA  
<213> unidentified

<220>  
<221> misc\_feature  
<222> (6)..(9)  
<223> any nucleotide

<400> 28  
cgaacnnng ttcg

14

<210> 29  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 29  
ctagacgttg acattgatta ttgac 25

<210> 30  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 30  
tctagagcca gtaagcagtg ggttc 25

<210> 31  
<211> 28  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 31  
gcggccgccc attgcattgg gattggtc 28

<210> 32  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 32  
gcggccgccc tccaagtctt tatctg 26

<210> 33  
<211> 28  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 33  
ggtaccggct cgtagagaaa tatgaacc 28

<210> 34  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 34  
ctcgagagag actgaatttc cctttt 26

<210> 35  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 35  
gcggccgcga aggaagatag ccaaggctta 30

<210> 36  
<211> 32  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 36  
gcggccgctt aaagatgcca cactgctggt ct 32

<210> 37  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 37  
gtcgacacct gtcggctact gggata 26

<210> 38  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 38  
gatatcgtgg gactctgttt ccagca 26

<210> 39  
<211> 6795  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA vector

<220>  
<221> misc\_feature  
<222> (892)..(892)  
<223> unknown nucleotide

<400> 39  
aagtccggca cctcgtgcac gcggatttcg gctccaacaa tgtcctgacg gacaatggcc 60  
gcataacagc ggtcattgac tggagcgagg cgatgttcgg ggattcccaa tacgaggtcg 120  
ccaacatctt cttctggagg ccgtgggttg cttgtatgga gcagcagacg cgctacttcg 180  
agcggaggca tccggagctt gcaggatcgc cgcggctccg ggcgtatatg ctccgcattg 240  
gtcttgacca actctatcag agcttggttg acggcaattt cgatgatgca gcttggggcg 300  
agggtcgatg cgacgcaatc gtccgatccg gagccgggac tgtcggggcg acacaaatcg 360  
cccgcagaag cgcggccgctc tggaccgatg gctgtgtaga agtactcgcc gatagtggaa 420  
accgacgccc cagcactcgt ccgagggcaa aggaatgcct gagaaaggaa gtgagctgta 480  
aaggctgagc tctctctctg acgtatgtag cctctgggta gcttcgtcac tcaactgtct 540  
tgactcagca tggcaatctg atgaaatccc agctgtaagt ctgcataaat tgatgatcta 600  
ttaacaata aagatgtcca ctaaaatgga agttttttac tgtcataactt tgttaagaag 660  
ggtgagaaca gagtacctac attttgaatg gaaggattgg agctacgggg gtgggggtgg 720  
gggtgggatt agataaatgc ctgctcttta ctgaaggctc tttactattg ctttatgata 780  
atgtttcata gttggatatc ataatttaaa caagcaaac caaattaagg gccagctcat 840  
tctccactc atgatctata gatctataga tctctcgtgg gatcattgtt tntctgatcc 900  
actggaagct tatcgatacc gtcgacctc agggggggcc cggtaaccag cttttgttcc 960  
ctttagttag ggttaattgc gcgcttggcg taatcatggt catagctgtt tctgtgtga 1020  
aattgttata cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc 1080  
tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgccttc 1140  
cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc 1200  
ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgtcgcg ctcggtcgtt 1260  
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 1320  
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 1380  
aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat 1440  
cgacgctcaa gtcagagggt gcgaaacccg acaggactat aaagatacca ggcgtttccc 1500  
cctggaagct cctcgtgcg ctctcctgtt ccgacctgc cgcttaccgg atacctgtcc 1560  
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 1620  
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac 1680  
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggttaagaca cgacttatcg 1740  
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 1800  
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggatatctg 1860

gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 1920  
accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaaa 1980  
ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaac 2040  
tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcaccta gatcctttta 2100  
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 2160  
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcattccata 2220  
gttgccctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 2280  
agtgcctgaa tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac 2340  
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 2400  
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 2460  
gttggttgcca ttgctacagg catcgtgggtg tcacgctcgt cgtttggtat ggcttcattc 2520  
agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttggtg caaaaaagcg 2580  
gttagctcct tcggtcctcc gatcgttgctc agaagtaagt tggccgcagt gttatcactc 2640  
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 2700  
gtgactgggt agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 2760  
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc 2820  
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 2880  
agttcgatgt aaccactcgt tgcacccaac tgatcttcag catcttttac tttcaccagc 2940  
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aaggcgacaa 3000  
cggaaatggt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggg 3060  
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aatagggggt 3120  
ccgcgcacat ttccccgaaa agtgccacct aaattgtaag cgtaaatatt ttgttaaaat 3180  
tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa 3240  
tcccttataa atcaaaagaa tagaccgaga tagggttgag tggtgttcca gtttggaaca 3300  
agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg 3360  
gcgatggccc actacgtgaa ccatcaccct aatcaagttt tttggggctg aggtgccgta 3420  
aagcactaaa tcggaaccct aaaggaggcc cccgatttag agcttgacgg ggaaagccgg 3480  
cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa 3540  
gtgtagcggg cacgctgcgc gtaaccacca caccgcccgc gcttaatgcg ccgctacagg 3600  
gcgcgtccca ttcgccattc aggctgcgca actgttgagg agggcgatcg gtgcgggcct 3660  
cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa 3720  
cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgagcgc gcgtaatacg 3780

actcactata	gggcgaattg	gagctccacc	gcggtgcggc	cgggccatgc	aggccacgac	3840
atgataagat	acattgatga	gtttggacaa	accacaacta	gaatgcagtg	aaaaaaatgc	3900
tttattttgtg	aaattttgtga	tgctattgct	ttattttgtaa	ccattataag	ctgcaataaaa	3960
caagttcttg	tacagctcgt	ccatgccgag	agtgatcccc	gcggcgggtca	cgaactccag	4020
caggaccatg	tgatcgcgct	tctcgttggg	gtctttgctc	agggcgggact	ggtagctcag	4080
gtagtggttg	tcgggcagca	gcacggggcc	gtcgccgatg	ggggtgttct	gctggtagtg	4140
gtcggcgagc	tgacgctgc	cgtcctcgat	gttggtggcgg	atcttgaagt	tcaccttgat	4200
gccgttcttc	tgcttgtcgg	ccatgatata	gacgttggtg	ctgttgtagt	tgtactccag	4260
cttggtgccc	aggatggtgc	cgtcctcctt	gaagtcgatg	cccttcagct	cgatgcgggt	4320
caccaggggtg	tcgccctcga	acttcacctc	ggcgcgggtc	ttgtagttgc	cgtcgctcct	4380
gaagaagatg	gtgcgctcct	ggacgtagcc	ttcgggcatg	gcggacttga	agaagtcgtg	4440
ctgcttcatg	tggtcgggggt	agcggggcgaa	gcactgcagg	ccgtagccga	aggtggtcac	4500
gaggggtgggc	cagggcacgg	gcagcttgcc	ggtggtgcag	atgaacttca	gggtcagctt	4560
gccgtaggtg	gcacgcctc	cgccctcgcc	ggacacgctg	aacttggtggc	cgtttacgctc	4620
gccgtccagc	tcgaccagga	tgggcaccac	cccgggtgaac	agtcctcgc	ccttgctcac	4680
catggtggcg	ttagcttgat	ttgacagtgg	ctgggggttg	cgccgccggg	ttttatagga	4740
agccacagcg	gccactcgag	ccataaaagg	caactttagg	aacggcgggg	ggtgattgga	4800
ttcgagtcgt	ttattcaccg	gccttgccgc	acagtgcagc	atTTTTTTac	ccctctccc	4860
ctccttttgc	gggggaaaaa	aaaaaaaaaa	aaaaaaaaag	gagaagagaa	aaaaagcgag	4920
cgagagagaa	agcgagattg	aggaagagga	tgaagagttt	ggcgatgggt	gctgggtccg	4980
taggcccaga	tggacaagaa	tagccccgc	ccttgccggac	agtatcccat	tcagtgactc	5040
agatcagatc	aagcggccgc	tctagaacta	gtggatccgg	ataagccagt	aagcagtggtg	5100
ttctctagtt	agccagagag	ctctgcttat	atagacctcc	caccgtacac	gcctaccgcc	5160
catttgcgctc	aatggggcgg	agttgttacg	acattttgga	aagtcccgtt	gattttgggtg	5220
ccaaaacaaa	ctccattga	cgtcaatggg	gtggagactt	ggaaatcccc	gtgagtcaaa	5280
ccgctatcca	cgccattga	tgtactgcca	aaaccgcac	accatggtaa	tagcgatgac	5340
taatacgtag	atgtactgcc	aagtaggaaa	gtcccataag	gtcatgtact	gggcataatg	5400
ccaggcggggc	catttaccgt	cattgacgctc	aatagggggc	gtacttggca	tatgatacac	5460
ttgatgtact	gccaagtggg	cagtttaccg	taaatactcc	accattgac	gtcaatggaa	5520
agtccttatt	ggcgttacta	tgggaacata	cgtcattatt	gacgtcaatg	ggcgggggtc	5580
gttgggcgggt	cagccaggcg	ggccatttac	cgtaagttat	gtaacgcgga	actccatata	5640
tgggctatga	actaatgacc	ccgtaattga	ttactattaa	taactagtca	ataatcaatg	5700

tcaacccccgg gctgcaggaa ttctaccggg taggggagggc gcttttccca aggcagtctg 5760  
gagcatgcgc tttagcagcc ccgctggcac ttggcgcata acaagtggcc tctggcctcg 5820  
cacacattcc acatccaccg gtagcgccaa cgggctccct tctttggtgg ccccttcgcg 5880  
ccaccttcta ctctccct agtcaggaag tcccccccg ccccgagct cgctcgtgc 5940  
aggacgtgac aaatggaagt agcacgtctc actagtctcg tgcagatgga caagcaccgc 6000  
tgagcaatgg aagcgggtag gcctttgggg cagcggccaa tagcagcttg gtccttcgc 6060  
tttctgggct cagaggctgg gaaggggtgg gtccgggggc gggctcaggg gcgggctcag 6120  
gggcggggcg ggcgcgaagg tcctccggac ccggcattct gcacgcttca aaagcgcacg 6180  
tctgccgcgc tgttctctc ttctcatct ccgggccttc gacctgcatg aaaaagcctg 6240  
aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc gtctccgacc 6300  
tgatgcagct ctccggaggc gaagaatctc gtgctttcag cttcgatgta ggaggcgctg 6360  
gatatgtcct gcgggtaa at agctgcgccg atggtttcta caaagatcgt tatgtttatc 6420  
ggcactttgc atcgcccgcg ctcccgattc cggaagtgt tgacattggg gaattcagcg 6480  
agagcctgac ctattgcata tcccgccgtg cacagggtgt cacgttgcaa gacctgcctg 6540  
aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg atcgctgcgg 6600  
ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc ggtcaataca 6660  
ctacatggcg tgatttcata tgccgattg ctgatccca tgtgtatcac tggcaaactg 6720  
tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg atgctttggg 6780  
ccgaggactg ccccg 6795

<210> 40  
<211> 5821  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA vector

<400> 40  
ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaac 60  
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 120  
aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt 180  
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 240  
gttgccctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 300  
agtgtgcaa tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac 360  
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 420  
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 480



gttggtgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc	540
agctccgggtt cccaacgata aaggcgagtt acatgatccc ccatgtttgtg caaaaaagcg	600
gttagctcct tcggtcctcc gatcgtttgtc agaagtaagt tggccgcagt gttatcactc	660
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct	720
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc	780
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgtctc	840
atcattggaa aacgtttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc	900
agttcgatgt aaccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc	960
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca	1020
cggaaatggt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcaggg	1080
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt	1140
ccgcgcacat tccccgaaa agtgccacct gacgtcgacg gatcgggaga tctcccgatc	1200
ccctatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc cagtatctgc	1260
tccctgcttg tgtgttgagg gtcgctgagt agtgcgcgag caaaatttaa gctacaacaa	1320
ggcaaggctt gaccgacaat tgcataga atctgcttag ggtaggcgt tttgcgctgc	1380
ttcgcgatgt acggggccaga tatacgcggt gacattgatt attgactagt tattaatagt	1440
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta	1500
cggtaaatgg cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga	1560
cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt	1620
tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta	1680
ttgacgtcaa tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg	1740
actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt	1800
tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc	1860
acccatttga cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat	1920
gtcgttaaaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct	1980
atataagcag agctctctgg ctaactagag aaccactgc ttactggctt atcgaaatta	2040
atacgactca ctataggag acccaagctg gctagcgttt aaacttaagc ttgccaccat	2100
gaagctactg tcttctatcg aacaagcatg cgatatattg cgacttaaaa agctcaagtg	2160
ctccaaagaa aaaccgaagt gcgccaagt tctgaagaac aactgggagt gtcgctactc	2220
tcccaaaacc aaaaggtctc cgctgactag ggcacatctg acagaagtgg aatcaaggct	2280
agaaagactg gaacagctat ttctactgat ttttctcga gaagacctg acatgatttt	2340
gaaaatggat tctttacagg atataaaagc attgttaaca ggattatttg tacaagataa	2400

tgtgaataaa	gatgccgtca	cagatagatt	ggcttcagtg	gagactgata	tgcctctaac	2460
attgagacag	ttgactgtat	cgccgggtac	cgagctcgga	tccactagtc	cagtgtggtg	2520
gaattctgca	gatatccagc	acagtggcgg	ccgctcgagt	ctagagggcc	cgtttaaacc	2580
cgctgatcag	cctcgactgt	gccttctagt	tgccagccat	ctgttgtttg	cccctcccc	2640
gtgccttcct	tgaccctgga	aggtgccact	cccactgtcc	tttcctaata	aaatgaggaa	2700
attgcatcgc	attgtctgag	taggtgtcat	tctattctgg	ggggtggggt	ggggcaggac	2760
agcaaggggg	aggattggga	agacaatagc	aggcatgctg	gggatgcggt	gggctctatg	2820
gcttctgagg	cggaaagaac	cagctggggc	tctagggggt	atccccacgc	gccctgtagc	2880
ggcgcattaa	gcgcggcggg	tgtggtggtt	acgcgcagcg	tgaccgctac	acttgccagc	2940
gccctagcgc	ccgctccttt	cgctttcttc	ccttcctttc	tcgccacgtt	cgccggcttt	3000
ccccgtcaag	ctctaaatcg	ggggctccct	ttagggttcc	gatttagtgc	tttacggcac	3060
ctcgacccca	aaaaacttga	ttagggtgat	ggttcacgta	gtgggccatc	gccctgatag	3120
acggtttttc	gccctttgac	gttgaggtcc	acgttcttta	atagtggact	cttgttccaa	3180
actggaacaa	cactcaacc	tatctcggtc	tattcttttg	atttataagg	gattttgccg	3240
atctcggcct	attgggttaa	aaatgagctg	atttaacaaa	aatttaacgc	gaattaattc	3300
tgtggaatgt	gtgtcagtta	gggtgtggaa	agtccccagg	ctccccagca	ggcagaagta	3360
tgcaaagcat	gcatctcaat	tagtcagcaa	ccagggtgtg	aaagtcccca	ggctccccag	3420
caggcagaag	tatgcaaagc	atgcatctca	attagtcagc	aaccatagtc	ccgcccctaa	3480
ctccgcccac	cccgccccta	actccgcca	gttccgcca	ttctccgccc	catggctgac	3540
taattttttt	tatttatgca	gaggccgagg	ccgcctctgc	ctctgagcta	ttccagaagt	3600
agtgaggagg	cttttttgga	ggcctaggct	tttgcaaaaa	gctcccggga	gcttgatat	3660
ccatttttcg	atctgatcaa	gagacaggat	gaggatcggt	tcgcatgatt	gaacaagatg	3720
gattgcacgc	aggttctccg	gccgcttggg	tggagaggct	attcggctat	gactgggcac	3780
aacagacaat	cggtgctct	gatgccgccg	tgttcgggt	gtcagcgcag	gggcgcccgg	3840
ttctttttgt	caagaccgac	ctgtccggtg	ccctgaatga	actgcaggac	gaggcagcgc	3900
ggctatcgtg	gctggccacg	acgggcgttc	cttgccgcagc	tgtgctcgac	gttgctactg	3960
aagcgggaag	ggactggctg	ctattgggcg	aagtgccggg	gcaggatctc	ctgtcatctc	4020
accttgctcc	tgccgagaaa	gtatccatca	tggctgatgc	aatgcggcgg	ctgcatacgc	4080
ttgatccggc	tacctgcca	ttcgaccacc	aagcgaacaa	tcgcatcgag	cgagcacgta	4140
ctcggatgga	agccggtctt	gtcgatcagg	atgatctgga	cgaagagcat	caggggctcg	4200
cgccagccga	actgttcgcc	aggctcaagg	cgcgcagtc	cgacggcgag	gatctcgtcg	4260
tgacccatgg	cgatgcctgc	ttgccgaata	tcatgggtga	aaatggccgc	ttttctggat	4320

```

tcacgcactg tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc 4380
gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta 4440
tcgccgctcc cgattcgcag cgcacgcctt tctatcgcct tcttgacgag ttcttctgag 4500
cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat cacgagattt 4560
cgattccacc gccgccttct atgaaagggt gggcttcgga atcgttttcc gggacgccgg 4620
ctggatgata ctccagcgcg gggatctcat gctggagttc ttcgcccacc ccaacttggt 4680
tattgcagct tataatgggt acaaataaag caatagcatc acaaatttca caaataaagc 4740
atTTTTTTTca ctgcattcta gttgtgggtt gtccaaactc atcaatgtat cttatcatgt 4800
ctgtataaccg tcgacctcta gctagagctt ggcgtaatca tggtcatagc tgtttctcgt 4860
gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa 4920
agcctggggg gcctaattgag tgagctaact cacattaatt gcgttgcgct cactgcccgc 4980
tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag 5040
aggcggtttg cgtattgggc gctcttccgc ttctcgtc actgactcgc tgcgctcggt 5100
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 5160
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 5220
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 5280
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 5340
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 5400
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 5460
cagttcgggtg taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 5520
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt 5580
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 5640
tacagagttc ttgaagtggg ggctaacta cggctacact agaagaacag tatttggtat 5700
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 5760
acaaaccacc gctggtagcg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 5820
a 5821

```

```

<210> 41
<211> 10075
<212> DNA
<213> Artificial

```

```

<220>
<223> recombinant DNA molecule

```

```

<220>
<221> misc_feature

```

<222> (9040)..(9040)  
<223> unsequenced nucleotide

<400> 41  
ctggtagcgg tgggtttttt gtttgcaagc agcagattac ggcagaaaa aaaggatctc 60  
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 120  
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 180  
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 240  
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagttgcct 300  
gactccccgt cgtgtagata actacgatac gggagggcct accatctggc cccagtgtgt 360  
caatgatacc gcgagacca cgctcaccgg ctccagattt atcagcaata aaccagccag 420  
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 480  
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgccg aacgttggtg 540  
ccattgctac aggcacgtgt gtgtcacgct cgtcgttttg tatggcttca ttcagctccg 600  
gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct 660  
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtggtatca ctcatggtta 720  
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 780  
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 840  
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 900  
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 960  
tgtaaccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 1020  
ggtgagcaaa aacaggaagg caaatgccg caaaaaaggg aataagggcg acacggaaat 1080  
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 1140  
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 1200  
catttccccg aaaagtgcc cctaaattgt aagcgtaaat attttgttaa aattcgcgtt 1260  
aaatttttgt taaatcagct cattttttta ccaataggcc gaaatcggca aaatccctta 1320  
taaataaaaa gaatagaccg agatagggtt gagtgttggt ccagtttgga acaagagtcc 1380  
actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg 1440  
cccactacgt gaaccatcac cctaatacaag ttttttgggg tcgaggtgcc gtaaagcact 1500  
aaatcggaac cctaaaggga gccccgatt tagagcttga cggggaaagc cggcgaacgt 1560  
ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc 1620  
ggtcacgctg cgcgtaacca ccacaccgc cgcgcttaat gcgccgtac agggcgcgtc 1680  
ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct 1740  
attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg 1800

gttttcccag	tcacgacgtt	gtaaaacgac	ggccagtgag	cgcgcgtaat	acgactcact	1860
atagggcgaa	ttggagctcc	accgcggtgc	ggccggggcca	tgcaggccac	gacatgataa	1920
gatacattga	tgagtttggg	caaaccacaa	ctagaatgca	gtgaaaaaaaa	tgctttattt	1980
gtgaaatttg	tgatgctatt	gctttatttg	taaccattat	aagctgcaat	aaacaagttc	2040
ttgtacagct	cgtccatgcc	gagagtgatc	ccggcggcgg	tcacgaactc	cagcaggacc	2100
atgtgatcgc	gcttctcggt	gggggtctttg	ctcagggcgg	actggtagct	caggtagtgg	2160
ttgtcgggca	gcagcacggg	gccgtcgccg	atgggggtgt	tctgctggta	gtggctcggcg	2220
agctgcacgc	tgccgtcctc	gatgttgtgg	cggatcttga	agttcacctt	gatgccgttc	2280
ttctgcttgt	cggccatgat	atagacgttg	tggctgttgt	agttgtactc	cagcttgtgc	2340
cccaggatgt	tgccgtcctc	cttgaagtcg	atgcccttca	gctcgatgcg	gttcaccagg	2400
gtgtcgccct	cgaacttcac	ctcggcgcgg	gtctttagtg	tgccgtcgtc	cttgaagaag	2460
atgggtgcgt	cctggacgta	gccttcgggc	atggcggact	tgaagaagtc	gtgctgcttc	2520
atgtggtcgg	ggtagcgggc	gaagcactgc	aggccgtagc	cgaagggtgg	cacgaggggtg	2580
ggccagggca	cgggcagctt	gccggtgggtg	cagatgaact	tcagggtcag	cttgccgtag	2640
gtggcatcgc	cctcgccctc	gccggacacg	ctgaacttgt	ggccgtttac	gtcgccgtcc	2700
agctcgacca	ggatgggcac	caccccggtg	aacagctcct	cgcccttgct	caccatgggtg	2760
gcgttagctt	gatttgacag	tggctggggg	ttgcgccgcc	gggttttata	ggaagccaca	2820
gcggccactc	gagccataaa	aggcaacttt	aggaacggcg	gggggtgatt	ggattcgagt	2880
cgtttatcca	cgggccttgc	cgcacagtgc	agcatttttt	tacccctctc	cccctccttt	2940
tgccgggggaa	aaaaaaaaaa	aaaaaaaaaa	aaggagaaga	gaaaaaaagc	gagcgagaga	3000
gaaagcgaga	ttgaggaaga	ggatgaagag	tttggcgatg	ggtgctgggt	ccgtagggccc	3060
agatggacaa	gaatagcccc	cgcccttgcg	gacagtatcc	cattcagtga	ctcagatcag	3120
atcaagcggc	cgccagtgtg	atggatatct	gcagaattcc	agcacactgg	cggccgttac	3180
tagtggatcc	ctctggagct	gcctagtgtg	ggtgtctgct	ggaaacttct	aaaatgtctc	3240
ggcctattta	attggagaat	tactgaaga	attaaaggac	tcagtaaaga	agaggtttgg	3300
tttgttcatg	aaagtaagta	gcttgtgaaa	tgctttattt	tcctgtgact	aaaattccac	3360
cttgtaattc	ccaataagga	ttgcttcaag	ctatttaata	caaattaaac	ctgtttcacc	3420
atgctcatag	gaaaatgtca	aagtatcttc	attcaccaac	tgagggtctg	ctcttctttg	3480
aaggtagctc	tgagatttgt	tttgtaatgg	agacaaagac	ttatttaact	gttggggagg	3540
gggaaggagg	acttaggcca	ggcaggcggg	gatgtggggg	acgaaccatg	aggggcaatt	3600
tggctgcctc	tgaaaaggtc	cagagctacc	ttaaaagaca	ggctttctat	aacaagtggg	3660
cctttctttt	cctttctttt	tttttttgag	acagagtctc	actctgttgc	ccaagctgga	3720

gtgcagtggt	gcgatctcag	ctcactgcaa	cctccagctc	ccaggttcaa	gcaattctcc	3780
tgcctcagcc	tcccaagtag	ctgggattac	aggcaaccgc	taccatgccc	agctaatttt	3840
ttgtattttt	aatagagatg	gggtttcacc	atgttggcca	gggtggtctc	aaactcctga	3900
cctcaagtga	tccaccacc	tcagcctccc	aaagtgctgg	gattacagac	atgagccacc	3960
gcgcccagac	acagctggct	ttgtcgccca	ggctggaagt	gcagtggcac	gatctcggct	4020
cactgcaagc	tctgcctcct	gggttcacgc	cattctcctg	cctcagcttc	ccaagtagct	4080
gagactacag	gcgcccgcga	ccacgcccgg	ctaatttttt	gtatttttag	tggagatggg	4140
gtttcaccat	gttagccagg	atggctcttga	tctcctgacc	tcgtgatccg	cctgtcttgg	4200
cctcccaaag	tgctgggatg	acaggcgtga	gccaccacgc	ctggcttaca	agtggccctt	4260
tctaagcccc	aagatgtcac	tgtcttttca	ctgaaggcgg	aaaaagaatg	taggcatggt	4320
tgtctgagag	ggagggagta	gttttaggag	aaaggaccaa	gagcacattt	tgagtgggaa	4380
gtaagtgggt	tccttttcta	gccttatgat	atggtttggc	tctgtgtccc	cacccaaatc	4440
tcatggtgaa	ctgtaatccc	cagtgggtga	ggtgaggcct	ggcgggaggg	actggatcat	4500
gggggtgatc	cttcatgaat	ggtttagcac	catcctcttg	gtgctactct	agtgatagcg	4560
ttctcacaag	atctgatggg	tgaaaagtgt	gtgtctcttg	tttttgctcc	tgccatgcaa	4620
gacgcctact	cccacttcac	cttctgctat	gagtcaaatt	ccctgaggcc	tccccagaag	4680
cagaagcttc	ctgtacagcc	ttgtgggcca	cgagccaatt	aaacctcttt	tctttataaa	4740
ttaccaggtc	tcagggattt	ctttatagca	gactaataca	cctcaaagta	gggatttatt	4800
ttattattct	atTTTTTTTT	tgagataggg	tctcgatctg	tcaccagggc	tggagtgcag	4860
tggcatgatc	tcagcttgct	gtaacctctg	cctcccgggt	tgaagcaatt	ctcatgccta	4920
agccacccaa	gtagctggaa	ttacagggtg	gtgccaccct	gctcagctaa	tttttgatc	4980
tttagtagag	ttggggtttc	tccatgttgg	ccaggctggg	atcaaactcc	tggactcaag	5040
caatctgccc	accttggcct	ccccagagtg	ccaggattac	aggcatgacc	catcttgcc	5100
ggtcagtttt	TTTTTTTTTT	aatccagtta	atgctgggtg	ccatgactca	cacctgtaat	5160
gccagtgctt	tgggaggccg	aggtgggaag	atcacttgag	gccaggagtt	tgagaccagc	5220
ctgggaaaca	tctcatctct	acaaaaaac	aaaaacaaaa	caaaacaaca	acaacatcaa	5280
acatagccag	gtgtggtggg	gcacacctgt	agtcccagct	ccttgggagg	ctgaggcagg	5340
aggattgctt	gagcccagga	gttcaaggct	gcagtgaact	atgattgcac	cactgcattc	5400
cagcctgggc	aacagagcaa	gaacctatct	ctatttaaga	aaaaatgctc	ctaagaaatc	5460
atcttagtac	ttggtaagct	attaagggtt	tttttaagtg	tattaattac	atagccagta	5520
taataactta	gttgctaaga	gctttgaggt	ggctccaagc	aagaagaaag	gagttgttgt	5580
taagctacag	ctcagagtct	cagaatgaac	tgtttccatc	tctgctgtcc	tttgcttaac	5640

tgtgggataa	atataat	ctcaacattc	aacacatgtg	tccttaatcg	tcctctacga	5700
cactggttcc	atgttgccaa	gagaacccag	aacagccgca	aacacctact	ggattgttac	5760
tggcctttcc	tgagaggatg	acccgggctt	tgaccttggt	catgttgaag	ttcttcaagt	5820
aggcctcgta	gattctcttg	gccagagatt	tgagatctgc	agtttcagaa	tcttctatgt	5880
catgttcaca	ggtaagaatt	tctgctttca	gttttgcttt	ctcagatctt	ggcattcgtc	5940
cggtagccgg	cgatacagtc	aactgtcttt	gacctttggt	actactctct	tccgatgatg	6000
atgtcgcact	tattctatgc	tgtctcaatg	ttagaggcat	atcagtctcc	actgaagcca	6060
atctatctgt	gacggcatct	ttattcacat	tatcttgtag	aaataatcct	gttaacaatg	6120
cttttatatc	ctgtaaagaa	tccattttca	aaatcatgtc	aaggctctct	cgaggaaaaa	6180
tcagtagaaa	tagctgttcc	agtctttcta	gccttgattc	cacttctgtc	agatgtgccc	6240
tagtcagcgg	agaccttttg	gttttgggag	agtagcgaca	ctcccagttg	ttcttcagac	6300
acttggcgca	cttcggtttt	tctttgggag	acttgagctt	tttaagtcgg	caaatatcgc	6360
atgcttggtc	gatagaagac	agtagcttca	tggtaggcaag	cttaagttta	aacgctagaa	6420
ctagtggatc	cggataagcc	agtaagcagt	gggttctcta	gttagccaga	gagctctgct	6480
tatatagacc	tcccaccgta	cacgcctacc	gcccat	gtcaatgggg	cggagttggt	6540
acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	tgacgtcaat	6600
ggggtggaga	cttggaatc	cccgtagatc	aaaccgctat	ccacgcccac	tgatgtactg	6660
ccaaaaccgc	atcaccatgg	taatagcgat	gactaatagc	tagatgtact	gccaagtagg	6720
aaagtcccat	aaggatcatg	actgggcata	atgccaggcg	ggccatttac	cgatcattgac	6780
gtcaataggg	ggcgtagctg	gcatatgata	cacttgatgt	actgccaagt	gggcagttta	6840
ccgtaaatac	tccaccatt	gacgtcaatg	gaaagtccct	attggcggtta	ctatgggaac	6900
atacgtcatt	attgacgtca	atgggcgggg	gtcggtgggc	ggtagccag	gcgggccatt	6960
taccgtaagt	tatgtaacgc	ggaactccat	atatgggcta	tgaactaatg	accccgtaat	7020
tgattactat	taataactag	tcaataatca	atgtcaaccc	cgggctgcag	gaattctacc	7080
gggtagggga	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	7140
cacttggcgc	atcacaagtg	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	7200
caaccggctc	ccttcttttg	tggccccctc	gcgccacctt	ctactcctcc	cctagtcagg	7260
aagttcccc	ccgccccgca	gctcgcgctg	tgtaggacgt	gacaaatgga	agtagcacgt	7320
ctcactagtc	tcgtgcagat	ggacaagcac	cgctgagcaa	tggaagcggg	taggcctttg	7380
gggcagcggc	caatagcagc	ttggctcctt	cgctttctgg	gctcagaggc	tgggaagggg	7440
tgggtccggg	ggcgggctca	ggggcgggct	caggggcggg	gcgggcgcga	aggtcctccg	7500
gacccggcat	tctgcacgct	tcaaaagcgc	acgtctgccg	cgctgttctc	ctcttctca	7560

tctccggggc	ttcgacctgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	7620
ttctgatcga	aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	7680
ctcgtgcttt	cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	7740
ccgatggttt	ctacaaagat	cgttatgttt	atcggcactt	tgcacccggc	gcgctcccga	7800
ttccggaagt	gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	7860
gtgcacaggg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	7920
cggtcgcgga	ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	7980
gccattcgg	accgcaagga	atcgggtcaat	acactacatg	gcgtgatttc	atatgcgcga	8040
ttgctgatcc	ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	8100
tgcgcaggc	tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	8160
tcgtgcacgc	ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	8220
tcattgactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	8280
tctggaggcc	gtggttggt	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	8340
cggagcttgc	aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	8400
tctatcagag	cttggttgac	ggcaatttcg	atgatgcagc	ttggggcgag	ggtcgatgcg	8460
acgcaatcgt	ccgatccgga	gccgggactg	tcgggcgtac	acaaatcgcc	cgcagaagcg	8520
cggccgtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccc	8580
gcactcgtcc	gagggcaaag	gaatgcctga	gaaaggaagt	gagctgtaaa	ggctgagctc	8640
tctctctgac	gtatgtagcc	tctggtttagc	ttcgtcactc	actgttcttg	actcagcatg	8700
gcaatctgat	gaaatcccag	ctgtaagtct	gcataaattg	atgatctatt	aaacaataaa	8760
gatgtccact	aaaatggaag	ttttttactg	tcatactttg	ttaagaaggg	tgagaacaga	8820
gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	gtgggattag	8880
ataaatgcct	gctctttact	gaaggctctt	tactattgct	ttatgataat	gtttcatagt	8940
tggatatcat	aatttaaaca	agcaaaacca	aattaagggc	cagctcattc	ctccactcat	9000
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctgatccac	tggaagctta	9060
tcgataccgt	cgacctcgag	ggggggcccc	gtaccagct	tttgttccct	ttagtgaggg	9120
ttaattgcgc	gcttggcgta	atcatgggtca	tagctgtttc	ctgtgtgaaa	ttgttatccg	9180
ctcacaattc	cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	gggtgcctaa	9240
tgagtgagct	aactcacatt	aattgcgttg	cgctcactgc	ccgctttcca	gtcgggaaac	9300
ctgtcgtgcc	agctgcatta	atgaatcggc	caacgcgcgg	ggagaggcgg	tttgcgattt	9360
gggcgctctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	9420
gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	9480



ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 9540  
ctggcggtttt tccataggct ccgccccctt gacgagcatc aaaaaaatcg acgctcaagt 9600  
cagaggtggc gaaacccgac aggactataa agataccagg cgtttcccc cttgaagctcc 9660  
ctcgtgcgct ctctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 9720  
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 9780  
gttcgctcca agctgggctg tgtgcacgaa cccccggtt agcccgaccg ctgcgcctta 9840  
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 9900  
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 9960  
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 10020  
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaac caccg 10075

<210> 42  
<211> 8971  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA vector

<400> 42  
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg 60  
aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaaagaa 120  
ccagctgggg ctctaggggg tatccccacg cgccctgtag cggcgatta agcgcggcgg 180  
gtgtgggtgg taccgcgacg gtgaccgcta cacttgccag cgccctagcg cccgctcctt 240  
tcgctttctt cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc 300  
gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg 360  
attagggtga tgggtcacgt agtgggcat cgccctgata gacggttttt cgccctttga 420  
cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc 480  
ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattgggtta 540  
aaaatgagct gatttaacaa aaatttaacg cgaattaatt ctgtggaatg tgtgtcagtt 600  
agggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca tgcattctca 660  
ttagtcagca accagggtgtg gaaagtcccc aggtcccca gcaggcagaa gtatgcaaag 720  
catgcatctc aattagtcag caaccatagt cccgccccta actccgcca tcccgccctt 780  
aactccgccc agttccgccc attctccgcc ccatggctga ctaatttttt ttatttatgc 840  
agaggccgag gccgcctctg cctctgagct attccagaag tagtgaggag gcttttttgg 900  
aggcctagga ttttgcaaaa agctcccggg agcttgata tccattttcg gatctgatca 960  
gcacgtgatg aaaaagcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa 1020

gttcgacagc	gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	1080
cttcgatgta	ggagggcgctg	gatatgtcct	gcgggtaa	agctgcgccg	atgggtttcta	1140
caaagatcgt	tatgtttatc	ggcactttgc	atcggccgcg	ctcccgattc	cggaagtgtc	1200
tgacattggg	gaattcagcg	agagcctgac	ctattgcatc	tcccgccgtg	cacagggtgt	1260
cacgttgcaa	gacctgcctg	aaaccgaact	gcccgcgtgt	ctgcagccgg	tcgcggaggc	1320
catggatgcg	atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	1380
gcaaggaatc	ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	1440
tgtgtatcac	tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	1500
cgatgagctg	atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	tgcacgcgga	1560
tttcggctcc	aacaatgtcc	tgacggacaa	tggccgcata	acagcgggtca	ttgactggag	1620
cgaggcgatg	ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	1680
gttggttgt	atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgcagg	1740
atcgccgcgg	ctccggggct	atatgtctcg	cattggtctt	gaccaactct	atcagagctt	1800
ggttgacggc	aatttcgatg	atgcagcttg	ggcgagggt	cgatgcgacg	caatcgtccg	1860
atccggagcc	gggactgtcg	ggcgtaacaa	aatcgccgc	agaagcgcg	ccgtctggac	1920
cgatggctgt	gtagaagtac	tcgccgatag	tggaaaccga	cgccccagca	ctcgtccgag	1980
ggcaaaggaa	tagcacgtgc	tacgagattt	cgattccacc	gccgccttct	atgaaagggt	2040
gggcttcgga	atcgttttcc	gggacgccgg	ctggatgatc	ctccagcgcg	gggatctcat	2100
gctggagttc	ttcgcccacc	ccaacttggt	tattgcagct	tataatgggt	acaaataaag	2160
caatagcatc	acaaatttca	caaataaagc	atTTTTTTca	ctgcattcta	gttgtgggtt	2220
gtccaaactc	atcaatgtat	cttatcatgt	ctgtataccg	tcgacctcta	gctagagctt	2280
ggcgtaatca	tggatcatag	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	2340
caacatacga	gccggaagca	taaagtgtaa	agcctgggg	gcctaagag	tgagctaa	2400
cacattaatt	gcgttgcgct	cactgccgc	tttccagtcg	ggaaacctgt	cgtgccagct	2460
gcattaatga	atcggccaac	gcgcggggag	aggcggtttg	cgtattgggc	gctcttccgc	2520
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	2580
ctcaaaggcg	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	2640
agcaaaaagg	cagcaaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgTTTTTTcca	2700
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	2760
cccgacagga	ctataaagat	accaggcggt	ttcccctgga	agctccctcg	tgcgctctcc	2820
tgttccgacc	ctgccgctta	ccgatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	2880
gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	taggtcggtc	gctccaagct	2940

gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	3000
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	3060
gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtggg	ggcctaacta	3120
cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	3180
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtttttttgt	3240
ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	3300
tacggggctc	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	3360
atcaaaaagg	atcttcacct	agatcctttt	aaattaaana	tgaagtttta	aatcaatcta	3420
aagtatatat	gagtaaactt	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	3480
ctcagcgatc	tgtctatttc	gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	3540
tacgatacgg	gagggcttac	catctggccc	cagtgcctga	atgataccgc	gagacccacg	3600
ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	3660
tggtcctgca	actttatccg	cctccatcca	gtctattaat	tgttgccggg	aagctagagt	3720
aagtagttcg	ccagttaata	gtttgcgcaa	cgttgttgcc	attgctacag	gcatcgtggg	3780
gtcacgctcg	tcgtttggtg	tggtttcatt	cagctccggg	tcccaacgat	caagggcaggt	3840
tacatgatcc	cccatgttgt	gcaaaaaagc	ggtagctcc	ttcgggtcctc	cgatcgttgt	3900
cagaagtaag	ttggccgcag	tggtatcact	catggttatg	gcagcactgc	ataattctct	3960
tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggg	gagtactcaa	ccaagtcatt	4020
ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccc	gcgtcaatac	gggataatac	4080
cgcgccacat	agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	4140
actctcaagg	atcttaccgc	tggtgagatc	cagttcgatg	taaccacttc	gtgcacccaa	4200
ctgatcttca	gcatctttta	ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaggca	4260
aatgcccga	aaaaagggaa	taagggcgac	acggaaatgt	tgaataactca	tactcttcct	4320
ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	4380
atgtatttag	aaaaataaac	aataggggt	tccgcgcaca	tttccccgaa	aagtgccacc	4440
tgacgtcgac	ggatcgggag	atctccccgat	cccctatggg	gcactctcag	tacaatctgc	4500
tctgatgccg	catagttaag	ccagtatctg	ctccctgctt	gtgtgttgga	ggtcgctgag	4560
tagtgcgcca	gcaaaaattta	agctacaaca	aggcaagggt	tgaccgacaa	ttgcatgaag	4620
aatctgctta	gggttagggc	ttttgcgctg	cttcgcgatg	tacggggccag	atatacgcg	4680
tgacattgat	tattgactag	ttattaatag	taatcaatta	cggggtcatt	agttcatagc	4740
ccatatatgg	agttccgcgt	tacataactt	acggtaaata	gccccgcctg	ctgaccgccc	4800
aacgaccccc	gcccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	4860

actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat	4920
caagtgtatc atatgccaaag tacgccccct attgacgtca atgacggtaa atggcccgcc	4980
tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta	5040
ttagtcatcg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag	5100
cggtttgact cacggggatt tccaagtctc cacccttctg acgtcaatgg gagtttgttt	5160
tggcaccaaa atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa	5220
atgggcggtg ggcgtgtacg gtgggaggtc tatataagca gagctctctg gctaactaga	5280
gaaccctactg cttactggct tatcgaaatt aatacgactc actataggga gacccaagct	5340
ggctagcggt taaacttaag cttgccacca tgaagctact gtcttctatc gaacaagcat	5400
gcgatatttg ccgacttaaa aagctcaagt gctccaaaga aaaaccgaag tgcgccaaagt	5460
gtctgaagaa caactgggag tgtcgctact ctcccaaaac caaaagggtc ccgctgacta	5520
gggcacatct gacagaagtg gaatcaaggc tagaaagact ggaacagcta tttctactga	5580
tttttctctg agaagacctt gacatgattt tgaaaatgga ttctttacag gatataaaag	5640
cattgttaac aggattatth gtacaagata atgtgaataa agatgccgtc acagatagat	5700
tggcttcagt ggagactgat atgcctctaa cattgagaca gcatagaata agtgcgacat	5760
catcatcgga agagagtagt aacaaagggtc aaagacagtt gactgtatcg ccgggtaccg	5820
cagtggggat gtctcataat ggtaagtaaa cagtcatcac catatacttt attattctca	5880
ttatagctgc cagaccagtg gacactaaag ccattgccaa aaatgtgtac agtttttcca	5940
ccaaatgcca gaatttagaa tattgcatgg cgataaaaca tttctctttt aggtcagtgt	6000
ttttaagtt ttattataga acctttctct ctgtgggttg gcactctgcca tgaggagaaa	6060
agagacttga aaaatctggg ggattatggg aaaaacttta gtgttagaac aaatataatg	6120
atcatcatga ttaattttta tattgcctca tgcattaaaa acagaattaa aatttacttt	6180
tgtgtatcag tagtgtcatt ctgctttaa tgtaagaaaa tcctgaagat gtctgggtct	6240
gaaactttaa atggcatatg agaatttaga ttcatttaat tttttttttt tttttttttt	6300
tgctgtttat tatgcaaaga cgagtattcg tttccttcac ttcaagggaac gtccatgagt	6360
tgtggaaagg gtaaaacaaa ttatttttat agttctgcaa aacttcaaag ccatagaaac	6420
ataagagaat aaaactttct acttttgaaa cctgtgtcta aagtcaagct tatgactttg	6480
ggtgtcaact taagggaat gatgtagaag acctgagaaa agacccatgt tttcatgatc	6540
tgtggaacta actcagtaag gtcctggaag acaaacgtgt ttgaaaaaat aattcttgg	6600
cactatttgg ttctgattgg ttgcctggct gacaaaataa attttaggtg taacttgata	6660
gtattatatg tgtctagtat taatcagtta ttttcaatgc actcaaggaa taaaaaatgc	6720
caacatgatc aaattaaatt tgaagtttca tttcatcagc tatectatth tttctttctt	6780

ttgttttcat ctctctcaaa actgatctag taagaaggca gtcttcatct atatgtactc	6840
taaagcatct gagaaaatta aatttttgct ctatcatgca gatacttggt tttcacagtg	6900
gaaaccatca agatatttga gagtcactca ttgattaata acctaccctt taaaaatagt	6960
tcacaaaatt ctatctgtat gattactaat gggtcagttc ctctaaaggc aagaatgtat	7020
tttgaataga gctatcagca gtactgaaca ggatggaaat ttggtccaac tttcctcaac	7080
aattgactac atctgctgta tacaactgta gtcattggatt gcaaaataat ttgctgggtc	7140
agttccatca gtttcctttt tatttcagac ctcataaaac ataagggcaa gcttccaagt	7200
caagggttgt attttaggtg atgcaatcaa gttgcaataa acttatttca tcttcatgaa	7260
gaaatcattg actttgaggt cattatcact gagacagcct aggaaagcat gccagttccg	7320
cttttaacac taatttttgc ctattaaaaa attatcactt cacacattca tgtaaaatat	7380
aattcctggt acttacaggc agtgcataca ttgatgtaaa atatatttcc tgttgcttac	7440
aggcattgca tacgtttact taaaatcaat ctgcaattgc ttattgttta taggcaaata	7500
ttatttaaac cccaacctat gtagtatgga tgggtataatt ttctcctccc ataaatgtca	7560
aagaaagcta tcataaactt gatagggtcat attttgactt catatgacaa tttattttat	7620
ttctagagag ggtcaatctt gatagggagt gtttgttcag ttctagcaaa agcaagaaaa	7680
tatatgaatt aagtgtcatg acatcagaga cttaggcagt tcaacactta aaaaattgca	7740
aacataacta agaaaacctc agaggttttt atgggagttt gtttgctgca cagaaaataa	7800
acatcagact aacaacaagg ttttgattag ttttgctgt tagctgtatg aagaaaacat	7860
ttctaaatac aaagttttct aggaagttga gtcactattt aaaatccatt gctttgacca	7920
atgtgaagcc aacaggggtt tcattcatat agaatgcctt tttatcaaaa gaggctctga	7980
gtgttatacc aaaatcagca cccagctgtt ttcagtatgc gcaagaaaac atgaacgtcc	8040
cccaaagcaa atgtactcca ttccgtgtta acctatttc tgctccctgg tctctgatta	8100
gtctggttga tgaatgtagc ttcccctttg gccatgcact cacagctttg aaggaggctt	8160
ctgaaggaac tggtagggaa ccctaggggc tcttgcta at gttgttggat agtcactgaa	8220
atcagtgatc accatgtcct ttcactttga aaattctatg attctgtata atcaaacaac	8280
ctaactcact ggatttttca cttttcaacc agctattccg tgtcttacag tgctctgggtc	8340
ttcacagata ggtaggcacc tggcatgtca gttgcctgaa tttgaaagtt ttcacctgta	8400
tgtttttggt cgataaaaaat aaaaatgtaa tttatatatc tgaatcaggt ctgtatgtta	8460
tgatcaattg ctacagcaatt tcgggcagtt ggtttgatgg ttatgtagta atgtagcctg	8520
agagcagaaa tacagagcct ctgggctaga gaaagtataa atggcatcct aggctatgta	8580
gggttacagc tcttcagaag gaactttcat tttcattgtg acacatcgtc tacatgttgt	8640
agaagaacat agtttcagaa ttcttccagt tagaaacata gtttctcaaa tattcacttt	8700

cagcattggg tagaaaaagt acccaggtga tggactgatt gaggatccac tagtaacggc 8760  
cgccagtgtg ctggaattct gcagatatcc atcacactgg cggccgctcg agtctagagg 8820  
gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt 8880  
ttgccccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta 8940  
ataaaatgag gaaattgcat cgcattgtct g 8971

<210> 43  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 43  
cgggtccccg gcgatacagt caactgtct 29

<210> 44  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 44  
taaagcttgc caccatgaag ctactgtctt 30

<210> 45  
<211> 10514  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA vector

<400> 45  
ctaaattgta agcgttaata ttttgttaaa attcgcgta aatttttggt aaatcagctc 60  
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 120  
gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga acgtggactc 180  
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 240  
ctaatacagt tttttggggt cgaggtgccc taaagcacta aatcggaacc ctaaaggagg 300  
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa 360  
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 420  
cacacccgcc gcgcttaatg cgccgctaca gggcgcgctc cattcgccat tcaggctgcy 480  
caactgttgg gaagggcgat cgggtcgggc ctcttcgcta ttacgccagc tggcgaaaagg 540  
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttccagc cagcagcttg 600

taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaaat tggagctcca	660
ccgcggtgcg gccggggccat gcaggccaga catgataaga tacattgatg agtttggaca	720
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc	780
tttatttcta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt	840
ccatgccgag agtgatcccc gcggcggtca cgaactccag caggaccatg tgatcgcgct	900
tctcgttggg gtctttgtct agggcggaact ggtagctcag gtagtggttg tcgggcagca	960
gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc	1020
cgtcctcgat gttgtggcgg atcttgaagt tcacctgat gccgttcttc tgcttgctcg	1080
ccatgatata gacgttgtgg ctgtttagt tgtactccag cttgtgcccc aggatgttgc	1140
cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga	1200
acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct	1260
ggacgtagce ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcgggggt	1320
agcggggcga gcaactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg	1380
gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct	1440
cgcctcgcgc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga	1500
tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat	1560
ttgacagtgg ctgggggttg cgccgccggg ttttatagga agccacagcg gccactcgag	1620
ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg	1680
gccttgccgc acagtgcagc atttttttac ccctctctcc ctccttttgc gggggaaaaa	1740
aaaaaaaaaa aaaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg	1800
aggaagagga tgaagagttt ggcgatgggt gctggttccg tagggccaga tggacaagaa	1860
tagccccgc ccttgccgac agtatccat tcagtactc agatcagtca agcggccgcg	1920
tctaacctag gagccacca cagacactat actataacca cactgaattt tctctctat	1980
ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc	2040
agacatggtg gcttgccgct gtaatcccag cactttggcc aagtcgggca gctcacttga	2100
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac	2160
aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc	2220
aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaaagatca tgccattgca	2280
ctccagcctg ggcaacagag cgagactccg tctcaaaca aacaaaaca aaacaaaaaa	2340
acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt	2400
tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca	2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg	2520

tatttatcat	tattattatc	acctttgatc	actctccctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggtatcaac	aattgcttat	tgaaggaata	aatgggtgaa	tgacaacttt	2640
gcaaagtttc	aatttttaat	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
cctttctagt	tctaatatct	tatgtctata	gtagttcgac	agttacttta	aatcagatat	2760
taaatgattt	tcattttaca	aagaacagac	ctgcctttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgaataa	gtacatgtga	aaatacttgg	caaagtaata	gtgaaaatag	2880
agctgatctc	acaggaatgt	tttgaaaatt	aaaacaccgt	atgtaaaata	ggtaatctaa	2940
tctgtgggca	catattttaa	ctctcaataa	atattaactt	tgattattta	aaaaatacct	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgat	tggttgccat	ttccgtcaaa	3060
atgagaaaat	tttcttctgc	actgaattct	tcttttaact	aaattttaag	aaaaaaacca	3120
atgattaaaa	agtagataag	aaaatgataa	tataaaacca	ttagactaaa	gaaaaactag	3180
gtttagatga	taaaatttga	aaacccatct	tactatttct	cagatctttg	gagggcaatt	3240
accattatca	agataggtgg	gagacttcat	actcttgaat	tgtttaatgt	tgctgagagt	3300
ttaaaagtct	aacttgaatt	ctcagctcta	atcatttcaa	agtaacatat	tcacatataa	3360
attatacctc	caaggcaata	aaaactccag	aaaatattaa	ctttaaaccc	acatattttc	3420
agatattata	catacaattt	tatttgttat	ttcctgaggc	atcctctggt	tgttatatga	3480
atccataata	aatgtagaa	gagtctgttg	atctgggggtg	agttcagttt	tctcctacac	3540
tggaagaaa	ataaaagggtg	ttaggggaaga	tgtgtataaa	gccttctttc	tatatgtgca	3600
ttcatctgga	gggacctaat	agcagaaatc	ataaaaaaaaa	aaaggcataa	atacagcagg	3660
aactatgagt	tctaagagag	tgtataattt	ctttctctgg	gtcctcaggc	aggccatatc	3720
acatttcact	gaattttcct	atctataaaa	tatgaataac	aaaatattcc	acttgtgtat	3780
ttcttaggaa	taaatggtag	tctggatatg	aagactattg	agcccatcca	tattgaaggt	3840
cactgaaagt	tcaagcttta	attataacta	atgaaagggt	ttcatatagt	atgactacta	3900
tgaatactat	attcatatac	tattctacat	gaattccacc	tgggtgaagtc	atctgagagt	3960
attcatgtgg	aggaggggatc	aacaaaactgt	ggcccatgga	ccaaatccat	tggattttat	4020
aaagttttat	tggaacacag	tcacactcat	ttgtttatgt	attgtctctg	ctacttttgc	4080
actaaaatgg	cagagttgag	tagctgttgt	agataccatg	tggccaaagc	cctttaaaga	4140
aaaaaagttt	gtcaaccctt	aaaatagaga	gtcaacaatc	tggacaattc	agaccccaaa	4200
catctctttc	tccctctaca	gcatcttttg	aagagtgtga	ataattatgt	caactctttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgct	ttcatataaa	4320
atactttttc	ttaaataaat	acaaaaatac	gtatttgtaa	aactaaatgt	ggataaagtt	4380
ttctttgtca	ataatgtaat	aaatattgag	aagggtgtgg	tttggacaat	catagcttac	4440



tggtttttca gctaattggca aaataaaaatt ccatcaaattg tcctgtgcat tatcacaatg	4500
tgtgtgtccta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt	4560
tcttgagaaa tttgtaaccc cagtgggtag gactttttctc tattcatcca catacgtgta	4620
ttcacctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct	4680
tcccagtttg agactgtctt ctcaaaaaga gggaaaaaat ggtcatcaat actcagttgg	4740
aaaaaatcca ttttatacac caaatcattt ctgctggcaa aactaggtgta gtaggactgg	4800
aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa	4860
tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca	4920
gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact	4980
cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt	5040
cgaggtcact tgtcgcaagt cagcaccttc actgtcttct ctagacggcg atacagtcaa	5100
ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg	5160
tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt	5220
attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatact gtaaagaatc	5280
cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag	5340
tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggg	5400
tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgact tcggtttttc	5460
tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag	5520
tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga	5580
gagctctgct tatatagacc tcccaccgta cagcctacc gccatttgc gtcaatgggg	5640
cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat	5700
tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgccccat	5760
tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact	5820
gccaaagtagg aaagtcccat aaggatcatg actgggcata atgccaggcg ggccatttac	5880
cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt	5940
gggcagttta ccgtaaatac tccaccatt gacgtcaatg gaaagtcctt attggcggtta	6000
ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag	6060
gcggggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg	6120
accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg	6180
cccgtaacat gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc	6240
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgcttttagca gccccgctgg	6300
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc	6360

caaccggctc ccttcttttg tggcccttc ggcacacctt ctactcctcc ctagtcagg	6420
aagttcccc cgcgccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt	6480
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg	6540
gggcagcggc caatagcagc ttggctcctt cgctttcttg gctcagaggc tgggaagggg	6600
tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg	6660
gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttctca	6720
tctccggggc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt	6780
ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat	6840
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg	6900
ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga	6960
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctccgcc	7020
gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc	7080
cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg	7140
gcccattcgg accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga	7200
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg	7260
tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc	7320
tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg	7380
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct	7440
tctggaggcc gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc	7500
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac	7560
tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg	7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg	7680
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc	7740
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa	7800
agatgtccac taaaatggaa gtttttcttg tcatactttg ttaagaaggg tgagaacaga	7860
gtacctacat tttgaatgga aggattggag ctacgggggt ggggggtggg tgggattaga	7920
taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt	7980
ggatatcata atttaaacia gcaaaaccaa attaagggcc agctcattcc tcccactcat	8040
gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt	8100
ggttctaagt actgtggttt ccaaagtgtg cagtttcata gcctgaagaa cgagatcagc	8160
agcctctgtt ccacatacac ttcattctca gtattgtttt gccaaagtct aattccatca	8220
gatttttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc	8280

gataccgctcg	acctcgaggg	ggggcccggt	accagctttt	tgttcccttt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	8400
cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	gtgcctaattg	8460
agtgagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	cgggaaacct	8520
gtcgtgccag	ctgcattaat	gaatcggcca	acgcgcgggg	agaggcgggt	tgcgtattgg	8580
gcgctcttcc	gcttctctgc	tactgaactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	8640
ggatatcagct	cactcaaagg	cggtaatatc	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	8760
ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	8820
gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	8940
gggaagcgtg	gcgctttctc	atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	9000
tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttatac	9060
cggtaactat	cgtcttgagt	ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	9180
gtggcctaac	tacggctaca	ctagaaggac	agtatttggg	atctgcgctc	tgctgaagcc	9240
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300
cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgate	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	gttaagggat	9420
tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttgggtc	gacagttacc	aatgcttaat	9540
cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgctcgttag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattggtg	9780
ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	9840
tacaggcatc	gtggtgtcac	gctcgtcggt	tggtatgggt	tcattcagct	ccggttccca	9900
acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcgggta	gctccttcgg	9960
tcctccgatac	gttgtcagaa	gtaagtgggc	cgcagtgtta	tcactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctgggtgagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgtc	10140
aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	10200

ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	10260
cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	10380
actcatactc	ttcctttttc	aatattattg	aagcatttat	cagggttatt	gtctcatgag	10440
cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	10500
ccgaaaagtg	ccac					10514

<210> 46  
 <211> 9626  
 <212> DNA  
 <213> Artificial

<220>  
 <223> recombinant DNA vector

<400> 46	
gacggatcgg	gagatctccc gatccccctat ggtcgaactct cagtacaatc tgctctgatg 60
ccgcatagtt	aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat	ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag	gcgtttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac	tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg	cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360
cccgccatt	gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca	atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc	aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta	catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600
tcgctattac	catggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg	atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg	ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg 780
gtaggcgtgt	acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg	gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900
gtttaaactt	aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattctgc 960
agatatccag	cacagtggct tgcggccgcg aaggaagata gccaaaggctt actgaggctg 1020
gtggagggag	ccactgctgg gctcaccatg gaccgccgga tgtggggggc ccacgtcttc 1080
tgcgtgttga	gcccgttacc gaccgtgagt agccagctga gaccctctgg cctctgccac 1140
tgggctccag	atttgggagc cacagggcca actggagcct ggcggagacc ctactgccct 1200
tctctgtctc	tagagtcctt ccctacgagc aggtggcagg ctgtggcttg gagagcccct 1260

gacactggggg	ctccctccct	catctccctc	tgcccagcaa	tctgtagtga	ctctgggtag	1320
tgctattaac	atgtccttcc	aggggctgag	ctgagctcac	cttcttggtg	cttccccag	1380
ggcctggctg	tgctctggg	ggccacagag	actctgtctg	ccacaagaca	gtgcctcaga	1440
attggtgagg	caaagtcagg	ctccctctt	agtccctgca	gtttgctctt	tctctcagct	1500
gtgcacctgt	gatgcttcca	cacgctcctg	tggttaagctt	tctggaccct	tccttagccc	1560
tatgactccc	tgagccactc	tccagtttgg	tggtttcctt	tgaagagggg	gtggccaggg	1620
aaggaatgga	acctaaggag	ttggtggaga	ccagctgctt	ctgctccagc	acttagggga	1680
tccatcatct	tctcacgatg	ctgagggctg	gagaggcttt	tggcaaccaa	attgtactcc	1740
atgggctaga	gcgctggagc	acggggcatc	aggagagagg	gcccttcctg	aatgggctgc	1800
agtttccag	taactttcca	ggagggttact	gcaagtgaca	gcagggacca	ggctgggagt	1860
cacaaggttc	tgggatgggtg	agggcctgaa	gcattcttcc	tccgctgcct	tattggaggg	1920
tgctgcccct	cccgagcacc	ttgctgcacc	tgctcctgag	tgcacaggct	ttgagcttga	1980
atcctggacg	catgaaacta	gccgtgtgaa	ctgcaagttc	ctcagccttc	ctgagctctg	2040
cttttatctc	ttttgttaaa	agagaaaata	tccgctggcc	cgggtcactg	tgcattgagt	2100
ggaatctgat	ttgtcaaggg	ctgggcatag	gggtgcaaca	caggaaactt	cgactgcccc	2160
cctgtctgcc	atctgtgtgc	actctacagc	aaaagcttta	tttttagcag	gggtacccaa	2220
gatctggtgg	tggccattca	gcagatgggtg	ctagtagcaa	agctagaact	caatcactgt	2280
ttgccagtgt	gggcatcatc	ccctgcagat	accaggaaga	gcagcaggca	tagatgcaag	2340
gcctgggggt	ccagggcgat	ggagcctccc	ctcaggggtg	gggtgggggtcc	acgatgtgtg	2400
tgggggggtg	actgggagcc	ccaggatgtg	tttcgccagt	cctaattcag	caggagcatt	2460
tggataaatc	ttagaacgat	ctatccttcc	actcccacca	cttccctccc	tagaatctct	2520
ccttacatct	ggccctatgc	ccacacatcc	tgagtcatta	gtggggcttc	tctgaatttc	2580
cctccttctc	ttagaggtgt	gggcactgtc	tgttcatect	tttggcatga	ggctgctctc	2640
ctggcatcct	ttgccctcag	ttgcttcccc	aaattggccc	tgaaaagttt	cccaggcagg	2700
gttggggctt	gccagactgt	gcagtaggca	gaaggatagg	cagagaccac	agctcttcca	2760
acagcccagg	gtcaatatca	ggctgatggg	gtgagaaagt	cccatagagg	cccctctgcc	2820
ctcctgtctc	tccttccccg	ctttgtgcgc	tcagtctggt	gggagatgta	ggggagacgc	2880
aacggggctc	tgccctgaga	tggacctgag	gttactgagt	cctcgggagg	aagccaagag	2940
ctgagtcctg	ggcaccacc	aaaataagac	ctggagcctg	ggtagtgggt	gacatctgca	3000
ccctggggcc	ttgcagggcc	aggcaggaga	tggcagagct	gcgtccaagc	catggacca	3060
tggggtcagc	agggcgggac	aggagtgact	tctcctgcca	gggtagcact	gggtgggggtg	3120
accctcctta	ctgggctggc	tctccagatt	ttctcctatc	ccacaaatct	gcggcagagc	3180

tctcgggtct	ctccacgtgg	tcacccacca	tctattgtcc	ccactctggt	ggtcaaagtc	3240
tctactctct	gaccccagcc	ctgggtcttac	atctgctaca	ctatttgctg	tgccctgcatt	3300
tctcccaaag	aacactagct	gttaattaaa	catttgctga	catggatcct	gctagacatt	3360
gaggggtgtg	gtgtgtgggg	tgggggtggg	ggccatgggg	agagaaacat	tagtcctgag	3420
ctgatacttc	ccaaatacac	cttttcccac	ttcccttgtc	cttgcacatg	ctgtgctttc	3480
tgttccagat	gctcttctcc	ccactgtgtg	tgtctggctc	taaaaaatgc	tcccttgctc	3540
ccagggtccag	ctcaaagtc	acctcccca	tccctgtgag	aggacattc	cctcccatgg	3600
tacctgtctt	gcttcttttt	gttttttttt	tgtttttctg	agatggagtc	tcactttatc	3660
gccaggctg	gagtacagt	gcacaatctc	agctcactgt	aacctccgcc	tcctgggttc	3720
aagtgattct	cctgcctcag	cctcctgagt	agctgggatt	acaggagcgt	gtcatcacac	3780
ccagctaatt	tttgtatttt	tgtagagata	gtgttttgcc	acattgggtca	ggctgggtgt	3840
taactcctga	cgtcagggtga	tccgcctacc	ttggcctccc	aaagtgtctg	gattacaggg	3900
atgagccact	gcgcccggcc	cctgtcctgc	ttctgatcac	agtttgcgca	gggtctatta	3960
ggatccagcc	cttctccctt	ccccagcccc	tcctctcctg	tgaaagcctg	ggggaacaga	4020
catctacact	ggttgacttg	gatgtgtgtt	ttatacaaaa	atgtgtatgc	tgtataccca	4080
gttcacagt	cagaaagtat	gttttgtaaa	cttggcttcc	cttgatgta	tctgtttggg	4140
ctctggcctg	ctgggcgctc	agttttgggt	ttccttcggt	ttggggcggt	gggagggaga	4200
agtgtagaa	attcctgctg	ttcacttcca	gtttcctgga	gaacttggtta	aaatacaaa	4260
tgctggatcc	accctagatc	tctacatatt	cctgcttcca	tcccaatgga	tctctagctg	4320
gttcttgagg	cacactttca	gaagtactgc	tagaagattc	ccaaaaaac	cccaacttgt	4380
ttccagcttc	cagagatcat	ttctcatggc	cactgtgtag	cacaagatct	gggaggctcc	4440
taaaactata	atctttcaag	gttgctcagc	catcaagggt	atgatccttt	catctacaca	4500
tccatgaagc	cacacatcca	tccttacaca	gctctctcac	tcatccacca	tgcacccatc	4560
catgcagcca	gtcatctgtc	tatccatcca	tccacctccc	cagccatata	tttgtctgtc	4620
tgtctgtata	ctcatccatt	ccactagcct	tccttcttcc	ttttctgcct	tcctgccttc	4680
ctgccttctc	tccttcttcc	ctttttttat	gcagccttca	cctcattcat	gtgctcactc	4740
atccatgtgc	tcattcattc	atccatcttc	tctctcactt	actaagctac	ccatttgctc	4800
atgtgttctc	gcctgtgtgc	accctaata	gcagagactg	agacaaaggt	cagaaattta	4860
aaatgggttc	cagtggggta	aatccaaagt	attggcagag	ctgcattcct	cctggagact	4920
gtagagaata	atccatttcc	ttgccttttc	cttgggttca	tgccttttct	ccatcttcag	4980
accagcagt	tggcatcttt	aagcggccgc	tcgagtctag	agggcccggt	taaacccgct	5040
gatcagcctc	gactgtgcct	tctagttgcc	agccatctgt	tgtttgcccc	tcccccgctc	5100

cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg	5160
catcgcattg tctgagtagg tgtcattcta ttctgggggg tgggggtgggg caggacagca	5220
aggggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt	5280
ctgaggcgga aagaaccagc tggggctcta gggggtatcc ccacgcgccc tgtagcggcg	5340
cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc	5400
tagcgcgccg tccttttcgct ttctttccctt cctttctcgc caggttcgcc ggctttcccc	5460
gtcaagctct aaatcggggc atcccttttag ggttccgatt tagtgcttta cggcacctcg	5520
accccaaaaa acttgattag ggtgatgggt cacgtagtgg gccatcgccc tgatagacgg	5580
tttttcgccc tttagcgttg gagtccacgt tctttaatag tggactcttg ttccaaactg	5640
gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt ttgggggattt	5700
cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg	5760
gaatgtgtgt cagttagggg gtggaaagtc cccaggctcc ccaggcaggc agaagtatgc	5820
aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag	5880
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc	5940
cgcccatccc gccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa	6000
ttttttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc cagaagtagt	6060
gaggaggctt ttttggaggc ctaggctttt gcaaaaagct cccgggagct tgtatatcca	6120
ttttcggatc tgatcagcac gtgatgaaaa agcctgaact caccgcgacg tctgtcgaga	6180
agtttctgat cgaaaagttc gacagcgtct ccgacctgat gcagctctcg gagggcgaag	6240
aatctcgtgc tttcagcttc gatgtaggag ggcgtggata tgtcctgcgg gtaaatagct	6300
gcgccgatgg tttctacaaa gatcgttatg tttatcgga ctttgcatcg gccgcgctcc	6360
cgattccgga agtgcttgac attggggaat tcagcgagag cctgacctat tgcattctccc	6420
gccgtgcaca ggggtgtcacg ttgcaagacc tgcttgaac cgaactgccc gctgttctgc	6480
agccggctgc ggaggccatg gatgcgatcg ctgcggccga tcttagccag acgagcgggt	6540
tcggccccatt cggaccgcaa ggaatcggtc aatacactac atggcgtgat ttcatatgcg	6600
cgattgctga tccccatgtg tatcactggc aaactgtgat ggacgacacc gtcagtgcgt	6660
ccgtcgcgca ggctctcgat gagctgatgc tttgggcccga ggactgcccc gaagtccggc	6720
acctcgtgca cgcggatttc ggctccaaca atgtcctgac ggacaatggc cgcataacag	6780
cggtcattga ctggagcgag gcgatgttcg gggattccca atacgaggtc gccaacatct	6840
tcttctggag gccgtgggtg gcttgatgg agcagcagac gcgctacttc gagcggaggc	6900
atccggagct tgcaggatcg ccgcggctcc gggcgatatat gctccgcatt ggtcttgacc	6960
aactctatca gagcttggtt gacggcaatt tcgatgatgc agcttgggcg cagggtcgat	7020

gcgacgcaat cgtccgatcc ggagccggga ctgtcgggcg tacacaaatc gcccgcagaa	7080
gcgcggccgt ctggaccgat ggctgtgtag aagtactcgc cgatagtgga aaccgacgcc	7140
ccagcactcg tccgagggca aaggaatagc acgtgctacg agatttcgat tccaccgccg	7200
ccttctatga aaggttgggc ttcggaatcg ttttccggga cgccggctgg atgatcctcc	7260
agcgcgggga tctcatgctg gagttcttcg cccaccccaa cttgtttatt gcagcttata	7320
atggttacaa ataaagcaat agcatcacia atttcacaaa taaagcattt ttttactgc	7380
attctagtgt tggtttgtcc aaactcatca atgtatctta tcatgtctgt ataccgtcga	7440
cctctagcta gagcttggcg taatcatggc catagctgtt tcctgtgtga aattgttatc	7500
cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct	7560
aatgagttag ctaactcaca ttaattgcgt tgcgctcact gcccgccttc cagtcgggaa	7620
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta	7680
ttgggcgctc ttccgccttc tcgctcactg actcgtcgcg ctcggtcgtt cggctgcggc	7740
gagcgggtatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg	7800
caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt	7860
tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa	7920
gtcagagggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct	7980
ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc	8040
cttcgggaag cgtggcgctt tctcaatgct cacgctgtag gtatctcagt tcgggtgtagg	8100
tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct	8160
tatccggtaa ctatcgtctt gagtccaacc cggtaaagaca cgacttatcg ccaactggcag	8220
cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga	8280
agtggtgggc taactacggc tacactagaa ggacagtatt tggatatctgc gctctgctga	8340
agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacia accaccgctg	8400
gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaaa ggatctcaag	8460
aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag	8520
ggatttttgt catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat	8580
gaagttttta atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct	8640
taatcagtga ggcacctatc tcagcgatct gtctattttc ttcattccata gttgcctgac	8700
tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa	8760
tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac cagccagccg	8820
gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt	8880
gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca	8940



```

ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt 9000
cccaacgatc aaggcgagtt acatgatccc ccatgttggtg caaaaaagcg gttagctcct 9060
tcggctcctcc gatcgttggtc agaagtaagt tggccgcagt gttatcactc atggttatgg 9120
cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactgggtg 9180
agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg 9240
cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa 9300
aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt 9360
aaccactcg tgcaccaaac tgatcttcag catcttttac tttcaccagc gtttctgggt 9420
gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aaggcgaca cggaaatgtt 9480
gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca 9540
tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat 9600
ttccccgaaa agtgccacct gacgtc 9626

```

```

<210> 47
<211> 22
<212> DNA
<213> Artificial

```

```

<220>
<223> primer

```

```

<400> 47
gagaaccac tgcttactgg ct 22

```

```

<210> 48
<211> 21
<212> DNA
<213> Artificial

```

```

<220>
<223> primer

```

```

<400> 48
gcattgctgt gcagtctgta a 21

```

```

<210> 49
<211> 7354
<212> DNA
<213> Artificial

```

```

<220>
<223> recombinant DNA vector

```

```

<400> 49
cgcgattgct gatccccatg tgtatcactg gcaaactgtg atggacgaca ccgtcagtgc 60
gtccgctcgcg caggctctcg atgagctgat gctttggggc gaggactgcc ccgaagtccg 120
gcacctcgtg cacgcggatt tcggctcaa caatgtcctg acggacaatg gccgcataac 180

```

agcgggtcatt gactggagcg aggcgatgtt cggggattcc caatacgagg tgcccaacat	240
cttcttcttg aggccgtggg tggcttgtat ggagcagcag acgcgctact tcgagcggag	300
gcatccggag cttgcaggat cgccgcgggt cggggcgat atgctccgca ttgggtcttga	360
ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg cgcagggtcg	420
atgcgacgca atcgtccgat ccggagccgg gactgtcggg cgtacacaaa tcgcccgcag	480
aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg gaaaccgacg	540
ccccagcact cgtccgaggg caaaggaata gagtatgcag aaattgatga tctattaaac	600
aataaagatg tccactaaaa tggaagtttt tctgtcata ctttgtaag aagggtgaga	660
acagagtacc tacattttga atggaaggat tggagctacg ggggtggggg tggggtggga	720
ttagataaat gcctgtctt tactgaaggc tctttactat tgctttatga taatgtttca	780
tagttggata tcataattta aacaagcaaa accaaattaa gggccagctc attcctccca	840
ctcatgatct atagatctat agatctctcg tgggatcatt gtttttctct tgattcccac	900
tttgtggttc taagtactgt ggtttccaaa tgtgtcagtt tcatagcctg aagaacgaga	960
tcagcagcct ctgttcaca tacacttcat tctcagtatt gttttgcaa gttctaattc	1020
catcagattt tagtggacat ctttattggt taatagatca tcaatttctg catactaagc	1080
ttatcgatac cgtcgacctc gagggggggc ccggtaccca gcttttggtc cctttagtga	1140
gggttaattg cgcgcttggc gtaatcatgg tcatagctgt ttctgtgtg aaattgttat	1200
ccgctcacia ttccacacia catacgagcc ggaagcataa agtgtaaagc ctgggggtgcc	1260
taatgagtga gctaactcac attaatgctg ttgcgctcac tgcccgttt ccagtcggga	1320
aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgct	1380
attgggcgt cttccgcttc ctgcgtcact gactcgctgc gctcggctgt tcggctgcgg	1440
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac	1500
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg	1560
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca	1620
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc cctggaagc	1680
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc	1740
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag	1800
gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc	1860
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca	1920
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg	1980
aagtgggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg	2040
aagccagtta ccttcggaaa aagagttggg agctcttgat ccggcaaaaa aaccaccgct	2100

ggtagcgggtg	gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	2160
gaagatcctt	tgatcttttc	tacgggggtct	gacgctcagt	ggaacgaaaa	ctcacgttaa	2220
gggatttttg	tcattgagatt	atcaaaaagg	atcttcacct	agatcctttt	aaattaaaaa	2280
tgaagtttta	aatcaatcta	aagtatatat	gagtaaactt	ggtctgacag	ttaccaatgc	2340
ttaatcagtg	aggcacctat	ctcagcgatc	tgtctatttc	gttcatccat	agttgcctga	2400
ctccccgtcg	tgtagataac	tacgatacgg	gagggccttac	catctggccc	cagtgcctga	2460
atgataccgc	gagaccacag	ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	2520
ggaagggccg	agcgcagaag	tggtcctgca	actttatccg	cctccatcca	gtctattaat	2580
tgttgccggg	aagctagagt	aagtagttcg	ccagttaata	gtttgcgcaa	cgttggtgcc	2640
attgctacag	gcacgtgggt	gtcacgctcg	tcgtttggta	tggcttcatt	cagctccggt	2700
tccaacgat	caaggcgagt	tacatgatcc	ccatgttgt	gcaaaaaagc	ggttagctcc	2760
ttcggtcctc	cgatcgttgt	cagaagtaag	ttggccgcag	tgttatcact	catgggttatg	2820
gcagcactgc	ataattctct	tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactgggt	2880
gagtactcaa	ccaagtcatt	ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccc	2940
gcgtcaatac	gggataatac	cgcgccacat	agcagaactt	taaaagtgct	catcattgga	3000
aaacgttctt	cggggcgaaa	actctcaagg	atcttaccgc	tgttgagatc	cagttcgatg	3060
taaccactc	gtgcacccaa	ctgatcttca	gcacttttta	ctttcaccag	cgtttctggg	3120
tgagcaaaaa	caggaaggca	aaatgccgca	aaaaagggaa	taagggcgac	acggaaatgt	3180
tgaataactca	tactcttcct	ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	3240
atgagcggat	acatatttga	atgtatttag	aaaaataaac	aaataggggt	tccgcgcaca	3300
tttccccgaa	aagtgccacc	taaattgtaa	gcgttaatat	tttggttaaaa	ttcgcgttaa	3360
atttttgtta	aatcagctca	ttttttaacc	aataggccga	aatcggcaaa	atcccttata	3420
aatcaaaaga	atagaccgag	atagggttga	gtgttggtcc	agtttggaac	aagagtccac	3480
tattaaagaa	cgtggactcc	aacgtcaaag	ggcgaaaaac	cgtctatcag	ggcgatggcc	3540
cactacgtga	accatcacc	taatcaagtt	ttttggggtc	gaggtgccgt	aaagcactaa	3600
atcggaaccc	taaagggagc	ccccgattta	gagcttgacg	gggaaagccg	gcgaacgtgg	3660
cgagaaagga	aggggaagaaa	gcgaaaggag	cgggcgctag	ggcgctggca	agtgtagcgg	3720
tcacgctgcg	cgtaaccacc	acaccgccc	cgcttaatgc	gccgctacag	ggcgcgctccc	3780
attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	ggtgcggggc	tcttcgctat	3840
tacgccagct	ggcgaaaggg	ggatgtgctg	caaggcgatt	aagttgggta	acgccaggggt	3900
tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgagcg	cgcgtaatac	gactcactat	3960
agggcggaatt	ggagctccac	cgcggtgcgg	ccggggccatg	caggccagac	atgataagat	4020

acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttattttgtg	4080
aaattttgtga tgctattgct ttattttgtaa ccattataag ctgcaataaa caagttgtcg	4140
agttacttgt acagctcgtc catgccgaga gtgatccccg cggcgggtcac gaactccagc	4200
aggaccatgt gatcgcgctt ctcgttgggg tctttgctca gggcggactg gtagctcagg	4260
tagtggttgt cgggcagcag cacggggccg tcgccgatgg ggggtgttctg ctggtagtgg	4320
tcggcgagct gcacgctgcc gtccctcgatg ttgtggcgga tcttgaagtt caccttgatg	4380
ccgttcttct gcttgctcggc catgatatag acgttgtggc tgtttagtgg gtactccagc	4440
ttgtgccccca ggatgttgcc gtccctccttg aagtcgatgc ccttcagctc gatgcggttc	4500
accaggggtg cgcctcga cttcacctcg gcgcgggtct tgtagttgcc gtcgtccttg	4560
aagaagatgg tgcgctcctg gacgtagcct tcgggcatgg cggacttgaa gaagtcgtgc	4620
tgcttcatgt ggtcggggta gcgggcgaag cactgcaggc cgtagccgaa ggtggtcacg	4680
agggtggggc agggcacggg cagcttgccg gtggtgcaga tgaacttcag ggtcagcttg	4740
ccgtaggtgg catcgccctc gccctcgccg gacacgctga acttggtggc gtttacgtcg	4800
ccgtccagct cgaccaggat gggcaccacc ccggtgaaca gctcctcgcc cttgctcacc	4860
atggtggcgt tagcttgatt tgacagtggc tgggggttgc gccgccgggt tttataggaa	4920
gccacagcgg cactcagac cataaaaggc aacttttagga acggcggggg gtgattggat	4980
tcgagtcgtt tattcacggc ccttgccgca cagtgcagca tttttttacc ccctctcccc	5040
tcctttttgcg ggggaaaaaa aaaaaaaaaa aaaaaaaagg agaagagaaa aaaagcgagc	5100
gagagagaaa gcgagattga ggaagaggat gaagagtttg gcgatgggtg ctggttccgt	5160
aggcccagat ggacaagaat agcccccgcc cttgcggaca gtatccatt cagtgactca	5220
gatcagtcaa gcggccgctc tagacggcga tacagtcaac tgtctttgac ctttgttact	5280
actctcttcc gatgatgatg tcgcacttat tctatgctgt ctcaatgtta gaggcataac	5340
agtctccact gaagccaatc tatctgtgac ggcattcttta ttcacattat cttgtacaaa	5400
taatcctgtt aacaatgctt ttatatcctg taaagaatcc attttcaaaa tcatgtcaag	5460
gtcttctcga ggaaaaatca gtagaaatag ctgttccagt ctttctagcc ttgattccac	5520
ttctgtcaga tgtgccctag tcagcggaga ctttttggtt ttgggagagt agcgacactc	5580
ccagttgttc ttcagacact tggcgcactt cggtttttct ttggagcact tgagcttttt	5640
aagtcggcaa atatcgcatg cttgttcgat agaagacagt agcttcatgg tggcggatcc	5700
cgataagcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagacct	5760
cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg	5820
gaaagtcccg ttgatttttg tgccaaaaca aactcccatt gacgtcaatg ggggtggagac	5880
ttggaaaatcc ccgtgagtca aaccgctatc cagcccatt gatgtactgc caaaaccgca	5940

```

tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata 6000
aggatcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg 6060
gcgtacttgg catatgatac acttgatgta ctgccaaagtg ggcagtttac cgtaaatact 6120
ccacccattg acgtcaatgg aaagtccta ttggcggttac tatgggaaca tacgtcatta 6180
ttgacgtcaa tgggcggggg tcgttggggc gtcagccagg cgggccattt accgtaagtt 6240
atgtaacgcg gaactccata tatgggctat gaactaatga ccccgtaatt gattactatt 6300
aataactagt caataatcaa tgtcaacgcg tataatctggc ccgtacatcg cgaagcagcg 6360
caaaacgcct aaccctaccg gggctgcagg aattctaccg ggtaggggag gcgcttttcc 6420
caaggcagtc tggagcatgc gcttttagcag ccccgctggc acttggcgca tcacaagtgg 6480
cctctggcct cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggt 6540
ggccccttcg cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag 6600
ctcgcgtcgt gcaggacgtg acaaattggaa gtagcacgtc tcactagtct cgtgcagatg 6660
gacaagcacc gctgagcaat ggaagcgggt aggccttttg ggagcgggc aatagcagct 6720
tggtctcttc gctttctggg ctccagaggct gggaaggggt ggggtccggg gcgggctcag 6780
gggcggggctc agggggcggg cgggcgcgaa ggtcctccgg acccggcatt ctgcacgctt 6840
caaaagcgca cgtctgccgc gctgttctcc tcttctcat ctccgggcct tcgacctgca 6900
tgaaaaagcc tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagtctgaca 6960
gcgtctccga cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg 7020
taggagggcg tggatatgtc ctgcgggtaa atagctgcgc cgatgggttc tacaagatc 7080
gttatgttta tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg 7140
gggaattcag cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc 7200
.aagacctgcc tgaaaccgaa ctgcccgtg ttctgcagcc ggtcgcggag gccatggatg 7260
cgatcgctgc ggccgatctt agccagacga gcgggttcgg ccatttcgga ccgcaaggaa 7320
tcgggtcaata cactacatgg cgtgatttca tatg 7354

```

<210> 50

<211> 10514

<212> DNA

<213> Artificial

<220>

<223> recombinant DNA vector

<400> 50

```

ctaaattgta agcgtaata ttttgtaaa attcgcgtta aatttttggt aaatcagctc 60
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 120
gataggggtg agtggtgttc cagtttgtaa caagagtcca ctattaaaga acgtggactc 180

```

caacgtcaaa	gggcgaaaaa	ccgtctatca	gggcgatggc	ccactacgtg	aaccatcacc	240
ctaatcaagt	tttttggggg	cgaggtgccg	taaagcacta	aatcggaacc	ctaaagggag	300
ccccgattt	agagcttgac	ggggaaagcc	ggcgaacgtg	gcgagaaagg	aagggaagaa	360
agcgaaagga	gcgggcgcta	gggcgctggc	aagtgtagcg	gtcacgctgc	gcgtaaccac	420
cacacccgcc	gcgcttaatg	cgccgctaca	gggcgcgctc	cattcgccat	tcaggctgcg	480
caactgttgg	gaagggcgat	cggtgcgggc	ctcttcgcta	ttacgccagc	tggcgaaagg	540
gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	ttttcccagt	cacgacgttg	600
taaaacgacg	gccagtgagc	gcgcgtaata	cgactcacta	tagggcgaat	tggagctcca	660
ccgcggtgcg	gccggggccat	gcaggccaga	catgataaga	tacattgatg	agtttggaca	720
aaccacaact	agaatgcagt	gaaaaaaatg	ctttatattgt	gaaatttgtg	atgctattgc	780
tttatttgta	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcgt	840
ccatgccgag	agtgatcccc	gcggcgggtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcgttggg	gtctttgtct	agggcggact	ggtagctcag	gtagtgggtg	tcgggcagca	960
gcacggggcc	gtcgccgatg	ggggtgttct	gctggtagtg	gtcggcgagc	tgcacgctgc	1020
cgctctcgat	gttggtggcg	atcttgaagt	tcaccttgat	gccgttcttc	tgcttgctcg	1080
ccatgatata	gacgttggtg	ctgttgtagt	tgtactccag	cttggtgccc	aggatgttgc	1140
cgctctcctt	gaagtcgatg	cccttcagct	cgatgcgggt	caccaggggtg	tcgccctcga	1200
acttcacctc	ggcgcgggtc	ttgtagttgc	cgtcgtcctt	gaagaagatg	gtgcgctcct	1260
ggacgtagcc	ttcgggcatg	gcggacttga	agaagtcgtg	ctgcttcatg	tggtcggggg	1320
agcggggcga	gcactgcagg	ccgtagccga	aggtggtcac	gaggggtggc	cagggcacgg	1380
gcagcttgcc	ggtggtgcag	atgaacttca	gggtcagctt	gccgtaggtg	gcacgcacct	1440
cgcctcgcgc	ggacacgctg	aacttggtgc	cgtttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccgggtgaac	agctcctcgc	ccttgctcac	catggtggcg	ttagcttgat	1560
ttgacagtgg	ctgggggttg	cgccgccggg	ttttatagga	agccacagcg	gccactcgag	1620
ccataaaagg	caacttttag	aacggcgggg	ggtgattgga	ttcgagtcgt	ttattcaccg	1680
gccttgccgc	acagtgcagc	atttttttac	cccctctccc	ctccttttgc	gggggaaaaa	1740
aaaaaaaaaa	aaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctggttccg	tagggccaga	tggacaagaa	1860
tagccccgcg	ccttgccggc	agtatcccat	tcagtgactc	agatcagtca	agcggccgcg	1920
tetaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
ctcaaataat	cctattcttg	ttcatgcctc	agtgcccttat	aacaagcaga	tttctcggcc	2040
agacatggtg	gcttgccgct	gtaatcccag	cactttggcc	aagtcgggca	gctcacttga	2100

ggtcaggagt	ttgaattcag	cctgtccaat	atgggtgaaac	cccatctcta	ctaaaaatac	2160
aaaaattagc	tgggcatggg	ggtgtgtgcc	tgtagtccca	gctactcagg	agactgaggg	2220
aggagaattg	cttgaactcg	ggaagcagag	gttgcagtga	gccaagatca	tgccattgca	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaaacaa	aacaaaacaa	aaacaaaaaa	2340
acaagcagat	ctctcaactt	ggaatgctgc	tgcttctcca	tgccaactgg	aaaacttctt	2400
tttcttttag	actcaactca	aatatgatcc	tccctgaagt	ctctccatct	ttcctagaca	2460
caattagcaa	ccagttcctc	tgtgttcaaa	cagcacttcg	tgtgtatgtg	tatgtatgtg	2520
tatttatcat	tattattatc	acctttgatc	actctccctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggtatcaac	aattgcttat	tgaaggaata	aatgggtgaa	tgacaacttt	2640
gcaaagtttc	aatttttaat	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
cctttctagt	tctaatatct	tatgtctata	gtagttcgac	agttacttta	aatcagatat	2760
taaatgattt	tcattttaca	aagaacagac	ctgcctttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgaataa	gtacatgtga	aaatacttgg	caaagtaata	gtgaaaatag	2880
agctgatctc	acaggaatgt	tttgaaaatt	aaaacaccgt	atgtaaaata	ggtaatctaa	2940
tctgtgggca	catattttaa	ctctcaataa	atattaactt	tgattattta	aaaaatacct	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgat	tggttgccat	ttccgtcaaa	3060
atgagaaaat	tttcttctgc	actgaattct	tcttttaact	aaattttaag	aaaaaaacca	3120
atgattaaaa	agtagataag	aaaatgataa	tataaaacca	ttagactaaa	gaaaaactag	3180
gtttgataga	taaaatttga	aaacccatct	tactatttct	cagatctttg	gagggcaatt	3240
accattatca	agataggtgg	gagacttcat	actcttgaat	tgtttaatgt	tgctgagagt	3300
ttaaaagtct	aacttgaatt	ctcagctcta	atcatttcaa	agtaacatat	tcacatataa	3360
attatacctc	caaggcaata	aaaactccag	aaaatattaa	ctttaaaccc	acatattttc	3420
agatattata	catacaattt	tatttgttat	ttcctgaggg	atcctctgtt	tgttatatga	3480
atccataata	aatgtagaa	gagtctgttg	atctgggggtg	agttcagttt	tctcctacac	3540
tggcaagaaa	ataaaagggtg	ttagggaaga	tgtgtataaa	gccttctttc	tatatgtgca	3600
ttcatctgga	gggacctaat	agcagaaatc	ataaaaaaaaa	aaaggcataa	atacagcagg	3660
aactatgagt	tctaagagag	tgtataattt	ctttctctgg	gtcctcaggg	aggccatatc	3720
acatttcact	gaattttcct	atctataaaa	tatgaataac	aaaatattcc	acttgtgtat	3780
ttcttaggaa	taaatggtag	tctggatatg	aagactattg	agccatttca	tattgaagggt	3840
cactgaaagt	tcaagcttta	attataacta	atgaaagggt	ttcatatagt	atgactacta	3900
tgaatactat	attcatatac	tattctacat	gaattccacc	tggtgaagtc	atctgagagt	3960
attcatgtgg	aggagggatc	aacaaactgt	ggcccatgga	ccaaatccat	tggattttat	4020

aaagttttat	tggaacacag	tcacactcat	ttgtttatgt	attgtctctg	ctacttttgc	4080
actaaaatgg	cagagttgag	tagctgttgt	agataccatg	tggccaaagc	ccttttaaaga	4140
aaaaaagttt	gtcaaccctt	aaaatagaga	gtcaacaatc	tggacaattc	agaccccaaa	4200
catctctttc	tccctctaca	gcctcttttg	aagagtgtga	ataattatgt	caactctttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgct	ttcatataaa	4320
atactttttc	ttaaataaat	acaaaaatac	gtattttgtg	aactaaatgt	ggataaagtt	4380
ttctttgtca	ataatgtaat	aaatattgag	aagggtgtgg	tttggacaat	catagcttac	4440
tggtttttca	gctaattggc	aaataaaaatt	ccatcaaattg	tcctgtgcat	tatcacaatg	4500
tgtgtgtcta	aacaccttgg	atgtgtttcc	cctattttgt	cctaagttgc	tcaatggatt	4560
tcttgagaaa	tttgtaacct	cagtgggtag	gactttttct	tattcatcca	catagctgta	4620
ttcacctgt	tttactcatt	tataaggatt	aatccacagc	ctctattgag	accatgttct	4680
tcccagtttg	agactgtctt	ctcaaaaaga	gggaaaaaat	ggcatcaat	actcagttgg	4740
aaaaaatcca	ttttatacac	caaatcattt	ctgctggcaa	acactaggta	gtaggactgg	4800
aatggagtca	gagagaagga	aaactttcac	tttttagttt	atgcattgct	actatttgaa	4860
tgtttacatc	aaagcatgta	ttaggaatcc	attttaatga	acaattggac	tagaaattca	4920
gctgtgcctg	tggtcttaat	ttgaaagcct	cctggccagc	tcactatgta	cctagaaact	4980
cagttcctgc	cagtcttggc	agacaccatt	gcattattat	accctgcatg	actttgttgt	5040
cgaggtcact	tgtcgcaagt	cacgaccttc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtctttga	cctttgttac	tactctcttc	cgatgatgat	gtcgcaacta	ttctatgctg	5160
tctcaatggt	agaggcatat	cagtctccac	tgaagccaat	ctatctgtga	cggcatcttt	5220
attcacatta	tcttgtacaa	ataatcctgt	taacaatgct	tttatatcct	gtaaagaatc	5280
cattttcaaa	atcatgtcaa	ggtcttctcg	aggaaaaatc	agtagaaata	gctgttccag	5340
tctttctagc	cttgattcca	cttctgtcag	atgtgcccta	gtcagcggag	accttttggg	5400
tttgggagag	tagcgacact	cccagttggt	cttcagacac	ttggcgcaact	tcggtttttc	5460
tttggagcac	ttgagctttt	taagtccgca	aatatcgcat	gcttgttcga	tagaagacag	5520
tagcttcatg	gtggcggatc	ccgataagcc	agtaagcagt	gggttctcta	gttagccaga	5580
gagctctgct	tatatagacc	tcccaccgta	cacgcctacc	gccattttgc	gtcaatgggg	5640
cggagttggt	acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	5700
tgacgtcaat	ggggtggaga	cttggaatc	cccgtgagtc	aaaccgctat	ccacgcccc	5760
tgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820
gccaaagtagg	aaagtcccat	aaggatcatg	actgggcata	atgccaggcg	ggccatttac	5880
cgtcattgac	gtcaataggg	ggcgtacttg	gcattatgata	cacttgatgt	actgccaaagt	5940



gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtcctt attggcggtta 6000  
ctatgggaac atacgtcatt attgacgtca atggggcgggg gtcgttgggc ggtcagccag 6060  
gcggggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg 6120  
accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg 6180  
cccgtagatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc 6240  
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgcttttagca gccccgctgg 6300  
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc 6360  
caaccggctc ctttcttttg tggccccctt gcgccacctt ctactcctcc cctagtcagg 6420  
aagttcccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt 6480  
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg 6540  
gggcagcggc caatagcagc ttggctcctt cgctttcttg gctcagaggc tgggaagggg 6600  
tgggtccggg ggcgggctca gggggcgggt caggggcggg gcgggcgcga aggtcctccg 6660  
gaccggcat tctgcacgt tcaaaagcgc acgtctgccg cgctgttctc ctcttctca 6720  
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt 6780  
ttctgatcga aaagtctgac agcgtctccg acctgatgca gctctcggag ggcgaagaat 6840  
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg 6900  
ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga 6960  
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc 7020  
gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc 7080  
cggtcgcgga ggccatggat gcgacgctg cgcccgatct tagccagacg agcgggttcg 7140  
gcccattcgg accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga 7200  
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg 7260  
tcgcgagggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc 7320  
tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg 7380  
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct 7440  
tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc 7500  
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggg cttgaccaac 7560  
tctatcagag cttgggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg 7620  
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg 7680  
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc 7740  
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa 7800  
agatgtccac taaaatggaa gtttttcttg tcatactttg ttaagaaggg tgagaacaga 7860

gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	tgggattaga	7920
taaatgcctg	ctctttactg	aaggctcttt	actattgctt	tatgataatg	tttcatagtt	7980
ggatatcata	atttaaacia	gcaaaaccaa	attaagggcc	agctcattcc	tcccactcat	8040
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctcttgatt	cccactttgt	8100
ggttctaagt	actgtgggtt	ccaaatgtgt	cagtttcata	gcctgaagaa	cgagatcagc	8160
agcctctgtt	ccacatacac	ttcattctca	gtattgtttt	gccaaagtct	aattccatca	8220
gatttttagtg	gacatcttta	ttgtttaata	gatcatcaat	ttctgcatac	taagcttata	8280
gataccgtcg	acctcgaggg	ggggcccggt	accagctttt	tgttcccttt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	8400
cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	gtgcctaatt	8460
agtgagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	cgaggaaacct	8520
gtcgtgccag	ctgcattaat	gaatcggcca	acgcgcgggg	agaggcggtt	tgcgtattgg	8580
gcgctcttcc	gcttctctcg	tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	8640
ggtatcagct	cactcaaagg	cggtaatatc	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	8760
ggcggttttt	cataggtctc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	8820
gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	8940
gggaagcgtg	gcgctttctc	atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	9000
tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttata	9060
cggttaactat	cgtcttgagt	ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	9180
gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	tgtgaagcc	9240
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300
cggtgggttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	gttaagggat	9420
tttggtcatt	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	9540
cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgctcgtgtg	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	9780

```

ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc 9840
tacaggcatc gtggtgtcac gtcgctcggt ttggtatggct tcattcagct ccggttccca 9900
acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcgggta gtccttcgg 9960
tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tctctcatgg ttatggcagc 10020
actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta 10080
ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc 10140
aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg 10200
ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc 10260
cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc 10320
aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat 10380
actcatactc ttcctttttc aatattattg aagcatttat caggggttatt gtctcatgag 10440
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc 10500
ccgaaaagtg ccac 10514

```

```

<210> 51
<211> 10247
<212> DNA
<213> Artificial

```

```

<220>
<223> recombinant DNA molecule

```

```

<220>
<221> misc_feature
<222> (2626)..(2626)
<223> unsequenced nucleotide

```

```

<400> 51
ctagaactag tggatccgga taagccagta agcagtgggt tctctagtta gccagagagc 60
tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgta atggggcgga 120
gttggttacga cattttggaa agtcccgttg attttggtgc caaaacaaac tccattgac 180
gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac gccattgat 240
gtactgcaa aaccgcatca ccatggtaat agcgatgact aatacgtaga tgtactgcca 300
agtaggaaag tcccataagg tcatgtactg ggcataatgc caggcgggcc atttaccgtc 360
attgacgtca atagggggcg tacttggcat atgatacact tgatgtactg ccaagtgggc 420
agtttaccgt aaatactcca ccattgacg tcaatggaaa gtccctattg gcgttactat 480
gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc agccaggcgg 540
gccatttacc gtaagttatg taacgcggaa ctccatatat gggctatgaa ctaatgaccc 600
cgtaattgat tactattaat aactagtcaa taatcaatgt caaccccggg ctgcaggaat 660

```

tctaccgggt aggggaggcg cttttcccaa ggcagtctgg agcatgcgct ttagcagccc	720
cgctggcact tggcgcatca caagtggcct ctggcctcgc acacattcca catccaccgg	780
tagcgccaac eggctccctt ctttggtggc cccttcgcgc caccttctac tcctccccta	840
gtcaggaagt tccccccgc cccgcagctc gcgtcgtgca ggacgtgaca aatggaagta	900
gcacgtctca ctagtctcgt gcagatggac aagcaccgct gagcaatgga agcgggtagg	960
cccttggggc agcggccaat agcagcttgg ctcttcgct ttctgggctc agaggctggg	1020
aaggggtggg tccggggcg ggctcagggg cgggctcagg ggcggggcg gcgcgaaggt	1080
cctccggacc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct gttctctct	1140
tcctcatctc cgggccttcg acctgcatga aaaagcctga actcaccgcg acgtctgtcg	1200
agaagtttct gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggaggggcg	1260
aagaatctcg tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata	1320
gctgcgccga tggtttctac aaagatcggt atgtttatcg gcactttgca tcggccgcgc	1380
tcccgattcc ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct	1440
cccgcctgac acaggggtgc acgttgcaag acctgcctga aaccgaactg cccgctgttc	1500
tgcagccggt cgcggaggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg	1560
ggttcggccc attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat	1620
gcgcgattgc tgatcccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg	1680
cgctcgtcgc gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc	1740
ggcacctcgt gcacgcggat ttgggtcca acaatgtcct gacggacaat ggccgcataa	1800
cagcggatcat tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca	1860
tcttctctct gagggcgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga	1920
ggcatccgga gcttgacgga tcgccgcggc tccgggcgta tatgctccgc attggtcttg	1980
accaactcta tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc	2040
gatgcgacgc aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgcga	2100
gaagcgcggc cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac	2160
gccccagcac tcgtccgagg gcaaaggaaat gcctgagaaa ggaagtgagc tgtaaaggct	2220
gagctctctc tctgacgtat gtagcctctg gttagcttcg tctactactg ttcttgactc	2280
agcatggcaa tctgatgaaa tcccagctgt aagtctgcat aaattgatga tctattaaac	2340
aataaagatg tccactaaaa tggaagtttt ttactgtcat actttgttaa gaagggtgag	2400
aacagagtac ctacattttg aatggaagga ttggagctac ggggggtggg gtgggggtgg	2460
gattagataa atgcctgctc ttactgaag gctctttact attgctttat gataatgttt	2520
catagttgga tatcataatt taaacaagca aaaccaaatt aagggccagc tcattcctcc	2580

actcatgatc tatagatcta tagatctctc gtgggatcat tgtttntctg atccactgga	2640
agcttatcga taccgtcgac ctcgaggggg ggcccgggtac ccagcttttg ttcccttttag	2700
tgaggggttaa ttgcgcgctt ggcgtaataca tgggtcatagc tgtttcctgt gtgaaattgt	2760
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctgggggt	2820
gcctaatagag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg	2880
ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag aggcgggttg	2940
cgtattgggc gctcttcgcg ttcctcgctc actgactcgc tgcgctcggc cgttcggctg	3000
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat	3060
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc	3120
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc	3180
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga	3240
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt	3300
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg	3360
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc	3420
gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt atcgccactg	3480
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc tacagagttc	3540
ttgaagtggg ggctaacta cggctacact agaaggacag tatttggtat ctgcgctctg	3600
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc	3660
gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgagaaa aaaaggatct	3720
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt	3780
taagggatth ttggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa	3840
aatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa	3900
tgttaataca gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc	3960
tgactccccg tcgtgtagat aactacgata cgggaggggt taccatctgg cccagtgct	4020
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca	4080
gccggaaggg ccgagcgag aagtggctct gcaactttat ccgcctccat ccagtctatt	4140
aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttggt	4200
gccattgcta caggcatcgt ggtgtcacgc tcgtcggttg gtatggcttc attcagctcc	4260
ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcgggttagc	4320
tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatgggt	4380
atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact	4440
ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc	4500

ccggcgtcaa	tacgggataa	taccgcgcca	catagcagaa	ctttaaaggt	gctcatcatt	4560
ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	atccagttcg	4620
atgtaacca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	cagcgtttct	4680
gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	4740
tgttgaatac	tcatactctt	cctttttcaa	tattattgaa	gcatttatca	gggttattgt	4800
ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	4860
acatttcccc	gaaaagtgcc	acctaaattg	taagcgttaa	tattttgtta	aaattcgcgt	4920
taaatTTTTg	ttaaatcagc	tcatttttta	accaataggc	cgaaatcggc	aaaatccctt	4980
ataaatcaaa	agaatagacc	gagatagggg	tgagtgttgt	tccagtttgg	aacaagagtc	5040
cactattaaa	gaacgtggac	tccaacgtca	aagggcgaaa	aaccgtctat	cagggcgatg	5100
gccactacg	tgaaccatca	ccctaataca	gttttttggg	gtcgaggtgc	cgtaaagcac	5160
taaatcgga	ccctaaaggg	agccccgat	ttagagcttg	acggggaaag	ccggcgaacg	5220
tggcgagaaa	ggaagggaag	aaagcgaaag	gagcgggcgc	tagggcgctg	gcaagtgtag	5280
cggtcacgct	gcgcgtaacc	accacacccg	ccgcgcttaa	tgcgccgcta	cagggcgcgct	5340
cccattcgcc	attcaggctg	cgcaactggt	gggaagggcg	atcgggtgcg	gcctcttcgc	5400
tattacgcca	gctggcgaaa	gggggatgtg	ctgcaaggcg	attaagttgg	gtaacgccag	5460
ggttttccca	gtcacgacgt	tgtaaaacga	cggccagtga	gcgcgcgtaa	tacgactcac	5520
tatagggcga	attggagctc	caccgcgggt	cggccggggc	atgcaggcca	cgacatgata	5580
agatacattg	atgagtttgg	acaaaccaca	actagaatgc	agtgaaaaaa	atgctttatt	5640
tgtgaaatTT	gtgatgctat	tgctttatTT	gtaaccatta	taagctgcaa	taaacaagtt	5700
cttgtacagc	tcgtccatgc	cgagagtgat	cccggcggcg	gtcacgaact	ccagcaggac	5760
catgtgatcg	cgcttctcgt	tggggctctt	gctcagggcg	gactggtagc	tcaggtagtg	5820
gttgctgggg	agcagcacgg	ggccgtcgcc	gatgggggtg	ttctgctggg	agtggctggc	5880
gagctgcacg	ctgccgtcct	cgatgttgtg	gcggatcttg	aagttcacct	tgatgccggt	5940
cttctgcttg	tcggccatga	tatagacggt	gtggctgttg	tagttgtact	ccagcttggt	6000
ccccaggatg	ttgccgtcct	ccttgaagtc	gatgcccttc	agctcgatgc	ggttcaccag	6060
gggtgctgccc	tcgaacttca	cctcggcgcg	ggctcttgtag	ttgccgtcgt	ccttgaagaa	6120
gatggtgcgc	tcctggacgt	agccttcggg	catggcggac	ttgaagaagt	cgtgctgctt	6180
catgtgggtc	gggtagcggg	cgaagcactg	caggccgtag	ccgaagggtg	tcacgagggg	6240
gggccagggc	acgggcagct	tgccggtggg	gcagatgaac	ttcaggggtc	gcttgccgta	6300
gggtggcatcg	ccctcgccct	cgccggacac	gctgaacttg	tggccgttta	cgctcgccgtc	6360
cagctcgacc	aggatggggc	ccaccccggt	gaacagctcc	tcgcccttgc	tcaccatggg	6420

ggcgttagct	tgatttgaca	gtggctgggg	gttgcgccgc	cgggttttat	aggaagccac	6480
agcggccact	cgagccataa	aaggcaactt	taggaacggc	ggggggtgat	tggattcgag	6540
tcgtttattc	accggccttg	ccgcacagtg	cagcattttt	ttaccccctc	tcccctcctt	6600
ttgcggggga	aaaaaaaaaa	aaaaaaaaaa	aaaggagaag	agaaaaaaag	cgagcgagag	6660
agaaagcgag	attgaggaag	aggatgaaga	gtttggcgat	gggtgctggt	tccgtaggcc	6720
cagatggaca	agaatagccc	ccgcccttgc	ggacagtatc	ccattcagtg	actcagatca	6780
gatcaagcgg	ccgccagtgt	gatggatata	tgcagaattc	cagcacactg	gcggccgtta	6840
ctagtggatc	ctcaatcagt	ccatcacctg	ggtacttttt	ctaccaatg	ctgaaagtga	6900
atatttgaga	aactatgttt	ctaactggaa	gaattctgaa	actatgttct	tctacaacat	6960
gtagacgatg	tgtcacaatg	aaaatgaaag	ttccttctga	agagctgtaa	ccctacatag	7020
cctaggatgc	catttatact	ttctctagcc	cagaggctct	gtatttctgc	tctcaggcta	7080
cattactaca	taaccatcaa	accaactgcc	cgaaattgct	gagcaattga	tcataacata	7140
cagacctgat	tcagatatat	aaattacatt	tttattttta	tcgtaccaa	acatacaggt	7200
gaaaactttc	aaattcaggc	aactgacatg	ccagggtgct	acctatctgt	gaagaccaga	7260
gcactgtaag	acacggaata	gctgggttgaa	aagtgaaaaa	tccagtgagt	tagggtgttt	7320
gattatacag	aatcatagaa	ttttcaaagt	gaaaggacat	ggtgatcact	gatttcagtg	7380
actatccaac	aacattagca	agagccccta	gggttccta	ccagttcctt	cagaagcctc	7440
cttcaaagct	gtgagtgc	ggccaaaggg	gaagctacat	tcatcaacca	gactaatcag	7500
agaccagggg	gcagaaataa	ggttaacacg	gaatggagta	catttgcttt	gggggacggt	7560
catgttttct	tgcgcatact	gaaaacagct	gggtgctgat	tttgggtata	cactcagagc	7620
ctcttttgat	aaaaaggcat	tctatatgaa	tgaaaaccct	gttggcttca	cattgggtcaa	7680
agcaatggat	tttaaatagt	gactcaactt	cctagaaaac	tttgtattta	gaaatgtttt	7740
cttcatacag	ctaacaggca	aaactaatca	aaaccttggt	gttagtctga	tgtttatttt	7800
ctgtgcagca	aacaaactcc	cataaaaacc	tctgagggtt	tcttagttat	gtttgcaatt	7860
ttttaagtgt	tgaactgcct	aagtctctga	tgtcatgaca	cttaattcat	atattttcct	7920
gcttttgcta	gaactgaaca	aacactccct	atcaagattg	accctctcta	gaaataaaat	7980
aaattgtcat	atgaagtcaa	aatatgacct	atcaagttta	tgatagcttt	ctttgacatt	8040
tatgggagga	gaaaattata	ccatccatac	tacatagggt	ggggttttaa	taatatttgc	8100
ctataaacia	taagcaattg	cagattgatt	ttaagtaaac	gtatgcaatg	cctgtaagca	8160
acaggaaata	tattttacat	caatgtatgc	actgcctgta	agtaacagga	attatatatt	8220
acatgaatgt	gtgaagtgat	aattttttta	taggcaaaaa	ttagtggtta	aagcggaact	8280
ggcatgcttt	cctaggctgt	ctcagtgata	atgacctcaa	agtcaatgat	ttcttcatga	8340

agatgaaata agttttattgc aacttgattg catcacctaa aatacaaccc ttgacttgga	8400
agcttgccct tatgttttat gaggtctgaa ataaaaagga aactgatgga actgaaccag	8460
caaattattht tgcaatccat gactacagtt gtatacagca gatgtagtca attggttgagg	8520
aaagttggac caaattttcca tctgtttcag tactgctgat agctctattc aaaatacatt	8580
cttgccctta gaggaactga cccattagta atcatacaga tagaattttg tgaactattht	8640
ttaaagggtg ggttattaat caatgagtga ctctcaaata tcttgatggt ttccactgtg	8700
aaaaccaagt atctgcatga tagagcaaaa atttaatttht ctcagatgct ttagagtaca	8760
tatagatgaa gactgccttc ttactagatc agttttgaga gagatgaaaa caaaagaaag	8820
aaaaaatagg atagctgatg aaatgaaact tcaaatttht tttgatcatg ttggcatttht	8880
ttattccctg agtgcattga aaataactga ttaatactag acacatataa tactatcaag	8940
ttacacctaa aattttattht gtcagccagg caaccaatca gaaccaaata gtgaccaaga	9000
attattthttht caaacacgtt tgtctttccag gaccttactg agttagttcc acagatcatg	9060
aaaacatggg tctttttctca ggtctttctac atcattttccc ttaagttgac acccaaagtc	9120
ataagcttga ctttagacac aggtttcaaaa agtagaaagt tttattctct tatgtttctca	9180
tggttttgaa gttttgcaga actataaaaa taatttgtht taccctttcc acaactcatg	9240
gacgttcctt gaagtgaagg aaacgaatac tcgtctttgc ataataaaca gcaaaaaaaaa	9300
aaaaaaaaaa aaaattaaat gaatctaaat tctcatatgc catttaaagt ttcagacca	9360
gacatcttca ggattttctt acattttaaag cagaatgaca ctactgatac acaaaagtaa	9420
attttaattc tgtttttaat gcatgaggca atataaaaa taatcatgat gatcattata	9480
tttgthtctaa cactaaagtt tttcccataa tccccagat ttttcaagtc tcttttctcc	9540
tcattggcaga tgcccaacca cagagagaaa ggttctataa taaaacttta aaaacactga	9600
cctaaaagag aaatgtttta tcgccatgca atattctaaa ttctggcatt tgggtggaaaa	9660
actgtacaca tttttggcaa tggcttttagt gtccactggt ctggcagcta taatgagaat	9720
aataaagtat atggtgatga ctgtttactt accattatga gacatcccca ctgcggtacc	9780
cggcgataca gtcaactgtc tttgaccttt gttactactc tcttccgatg atgatgtcgc	9840
acttattcta tgctgtctca atgttagagg catatcagtc tccactgaag ccaatctatc	9900
tgtgacggca tctttattca cattatcttg tacaataat cctgttaaca atgcttttat	9960
atcctgtaaa gaatccattt tcaaaatcat gtcaagggtc tctcgaggaa aaatcagtag	10020
aaatagctgt tccagtcttt ctagccttga ttccacttct gtcagatgtg ccctagtcag	10080
cggagacctt ttggtttttg gagagtagcg aactcccgag ttgttcttca gacacttggc	10140
gcacttcggt ttttcttttg agcacttgag ctttttaagt cggcaaatat cgcattgctg	10200
ttcgatagaa gacagtagct tcatggtggc aagcttaagt ttaaacg	10247



<210> 52  
<211> 10514  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA vector

<400> 52  
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aattttttgtt aaatcagctc 60  
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 120  
gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga acgtggactc 180  
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 240  
ctaatacagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaaggagag 300  
ccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa 360  
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 420  
cacaccgcc gcgcttaatg cgccgctaca gggcgctcc cattcgccat tcaggctgcg 480  
caactgttgg gaagggcgat cgggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540  
gggatgtgct gcaagggcat taagttgggt aacgccaggg ttttccagc cagcagcttg 600  
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca 660  
ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca 720  
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc 780  
tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt 840  
ccatgccgag agtgatcccg gcggcgggtc cgaactccag caggaccatg tgatcgcgct 900  
tctcgttggg gtctttgttc agggcggact ggtagctcag gtagtggttg tcgggcagca 960  
gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc 1020  
cgctctcgat gttgtggcgg atcttgaagt tcacctgat gccgttcttc tgcttgctcg 1080  
ccatgatata gacgttgttg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc 1140  
cgctctcctt gaagtcgatg cccttcagct cgatgcgggt caccagggtg tcgccctcga 1200  
acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct 1260  
ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt 1320  
agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gaggggtggc cagggcacgg 1380  
gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct 1440  
cgccctcgcc ggacacgctg aacttggtgc cgtttacgtc gccgtccagc tcgaccagga 1500  
tgggcaccac cccggtgaac agctcctcgc cttgtctcac catggtggcg ttagcttgat 1560  
ttgacagtgg ctgggggttg cgccgccggg ttttatagga agccacagcg gccactcgag 1620

ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg	1680
gccttgccgc acagtgcagc atttttttac cccctctccc ctcccttttg gggggaaaaa	1740
aaaaaaaaa aaaaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg	1800
aggaagagga tgaagagttt ggcgatgggt gctgggttccg taggcccaga tggacaagaa	1860
tagcccccgc ccttgcggaac agtatcccat tcagtgaactc agatcagtca agcggccgcg	1920
tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat	1980
ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc	2040
agacatggtg gcttgcgccct gtaatcccag cactttggcc aagtcgggca gctcacttga	2100
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac	2160
aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc	2220
aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaaagatca tgccattgca	2280
ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa	2340
acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt	2400
tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca	2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg	2520
tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt	2580
agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt	2640
gcaaagtttc aatttttaat ctattgattt gaaagaatta ttatacatta ttttaaatat	2700
cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat	2760
taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt	2820
agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag	2880
agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa	2940
tctgtgggca catatttaaa ctctcaataa atattaactt tgattattta aaaaatacct	3000
ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa	3060
atgagaaaaat tttcttctgc actgaattct tcttttaact aaattttaag aaaaaacca	3120
atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag	3180
gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt	3240
accattatca agatagggtg gagacttcat actcttgaat tgtttaatgt tgctgagagt	3300
ttaaaagtct aacttgaatt ctgagctcta atcatttcaa agtaacatat tcacatataa	3360
attatacctc caaggcaata aaaactccag aaaatattaa ctttaaacc acatattttc	3420
agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga	3480
atccataata aaatgtagaa gagtctgttg atctgggggtg agttcagttt tctcctacac	3540

tggcaagaaa ataaaagggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca	3600
ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg	3660
aactatgagt tctaagagag tgtataatct ctttctctgg gtcctcaggc aggccatctc	3720
acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat	3780
ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaagggt	3840
cactgaaagt tcaagcttta attataacta atgaaagggt ttcatatagt atgactacta	3900
tgaatactat attcatatac tattctacat gaattccacc tgggtgaagtc atctgagagt	3960
attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaattccat tggattttat	4020
aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctacttttgc	4080
actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cttttaaaga	4140
aaaaaagttt gtcaaccctt aaaatagaga gtcaacaatc tggacaattc agaccccaaa	4200
catctctttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg	4260
ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatatata	4320
atactttttc ttaaataaat acaaaaatac gtattttgtaa aactaaatgt ggataaagtt	4380
ttctttgtca ataattgtaaat aaatattgag aagggtgtgg tttggacaat catagcttac	4440
tgggtttttca gctaattggca aaataaaatt ccatcaaattg tcctgtgcat tatcacaatg	4500
tgtgtgtcta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt	4560
tcttgagaaa tttgtaacct cagtgggtag gacttttctc tattcatcca catacgtgta	4620
ttcaccctgt ttactcatt tataaggatt aatccacagc ctctattgag accatgttct	4680
tcccagtttg agactgtctt ctcaaaaaga gggaaaaaat ggatcatcaat actcagttgg	4740
aaaaaatcca ttttatacac caaatcattt ctgctggcaa aactaggta gtaggactgg	4800
aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa	4860
tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca	4920
gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact	4980
cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt	5040
cgaggtcact tgtcgcaagt cacgacctc actgtcttct ctagacggcg atacagtcaa	5100
ctgtctttga cttttgttac tactctcttc cgatgatgat gtgcgactta ttctatgctg	5160
tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt	5220
attcacatta tcttgtaaaa ataactctgt taacaatgct tttatatcct gtaaagaatc	5280
cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag	5340
tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggt	5400
tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgact tcggtttttc	5460

tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag	5520
tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga	5580
gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatTTgc gtcaatgggg	5640
cggagtTgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat	5700
tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgccc	5760
tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact	5820
gccaaagtagg aaagtcccat aaggTcatgt actgggcata atgccaggcg ggccatttac	5880
cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaa	5940
gggcagttta ccgtaaatac tccaccatt gacgtcaatg gaaagtcctt attggcgTta	6000
ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag	6060
gcggggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg	6120
accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg	6180
cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc	6240
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgcttttagca gcccgcgtgg	6300
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc	6360
caaccggctc ccttcttttg tggccccttc gcgccacctt ctactcctcc cctagtcagg	6420
aagttcccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt	6480
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg	6540
gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg	6600
tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg	6660
gacccgcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttctca	6720
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt	6780
ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat	6840
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctcgc	6900
ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga	6960
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc	7020
gtgcacaggg tgTcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc	7080
cggtcgcgga ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg	7140
gccattcgg accgcaagga atcggtcaat aactacatg gcgtgatttc atatgcgcga	7200
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg	7260
tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc	7320
tcgtgcacgc ggatttcggc tccaacaatg tctgacgga caatggccgc ataacagcgg	7380

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct	7440
tctggaggcc gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc	7500
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac	7560
tctatcagag cttgggtgac ggcaatttcg atgatgcagc ttgggcgagc ggtcgatgcg	7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg	7680
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc	7740
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaata	7800
agatgtccac taaaatggaa gtttttctg tcatactttg ttaagaagg tgagaacaga	7860
gtacctacat tttgaatgga aggattggag ctacgggggt gggggtggg tgggattaga	7920
taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt	7980
ggatatcata atttaaacaa gcaaaaccaa attaagggc agctcattcc tcccactcat	8040
gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt	8100
ggttctaagt actgtggttt ccaaagtgt cagtttcata gcctgaagaa cgagatcagc	8160
agcctctggt ccacatacac ttcatctca gtattgtttt gccaagttct aattccatca	8220
gatttttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttata	8280
gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt	8340
aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct	8400
cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaattg	8460
agtgagctaa ctcacattaa ttgcgttgcg ctactgccc gctttccagt cgggaaacct	8520
gtcgtgccag ctgcattaat gaatcgcca acgcgcgggg agaggcggtt tgcgtattgg	8580
gcgctcttcc gcttctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc	8640
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg	8700
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct	8760
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca	8820
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct	8880
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc	8940
gggaagcgtg gcgctttctc atagctcacg ctgtagggtat ctcaagtccg tgtaggctcg	9000
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttata	9060
cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac tggcagcagc	9120
cactggtaac aggattagca gagcgaggta tgtaggcggg gctacagagt tcttgaagtg	9180
gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc	9240
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag	9300

cggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaaac	gaaaactcac	gttaagggat	9420
tttgggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttgggtct	gacagttacc	aatgcttaat	9540
cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgtcgtgtag	ataactacga	tacggggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	9780
ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	9840
tacaggcatc	gtggtgtcac	gctcgtcgtt	tggtagggct	tcattcagct	ccggttccca	9900
acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcgggta	gtccttcggg	9960
tcctccgatc	gttgtcagaa	gtaagttggc	cgcagtgtta	tcactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgtc	10140
aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	10200
ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	10260
cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	10380
actcatactc	ttcctttttc	aatattattg	aagcatttat	caggggttatt	gtctcatgag	10440
cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	gggggttccgc	gcacatttcc	10500
ccgaaaagtg	ccac					10514

<210> 53  
 <211> 168  
 <212> DNA  
 <213> Homo sapiens

<400> 53	
cggaggacag	tactccgctc ggaggacagt actcgcctcg gaggacagta ctccgctcga 60
ggacagtact	ccgctcggag gacagtactc cgatccgctcg actctagtcc gctcggagga 120
cagtactccg	ctcggaggac agtactccgc tcggaggaca gtactccg 168

<210> 54  
 <211> 22  
 <212> DNA  
 <213> Artificial

<220>  
 <223> primer

<400> 54  
gagaacccac tgcttactgg ct 22

<210> 55  
<211> 20  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 55  
caagatcgcc ctgcctttg 20

<210> 56  
<211> 33  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 56  
atatggtacc gcagtgggga tgtctcataa tgg 33

<210> 57  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 57  
aattggatcc tcaatcagtc catcacctgg 30

<210> 58  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 58  
cgtctagaga gtgtgtcctg tcagaagaac 30

<210> 59  
<211> 31  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 59  
atcgggccgc actcctgacc tcaggtgatc c 31

<210> 60  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 60  
cgtctagaga agacagtga ggtcgtgac 29

<210> 61  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 61  
gcgcggccgc gtctaacctt ggagccac 29

<210> 62  
<211> 28  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 62  
gctctagagg acgaatgcc a gatctga 28

<210> 63  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> primer

<400> 63  
caagcggccg ccagtgtgat ggatatctg 29

<210> 64  
<211> 5218  
<212> DNA  
<213> Artificial

<220>  
<223> recombinant DNA

<400> 64  
gccggagcgg tcgagttctg gaccgaccgg ctcgggttct cccgggactt cgtggaggac 60

gacttcgccg gtgtggtccg ggacgacgtg accctgttca tcagcgcggg ccaggaccag 120

gtgggtgccg acaacaccct ggccctgggtg tgggtgcgcg gcctggacga gctgtacgcc 180



gagtggtcgg aggtcgtgtc cacgaacttc cgggacgcct ccgggcccgc catgaccgag	240
atcggcgagc agccgtgggg gcgggagttc gccctgcgcg acccggccgg caactgcgtg	300
cacttcgtgg ccgaggagca ggactgacac gtgctacgag atttcgattc caccgccgcc	360
ttctatgaaa ggttgggctt cggaatcggt ttccgggacg ccggctggat gatcctccag	420
cgcggggatc tcatgctgga gttcttcgcc caccccaact tgtttattgc agcttataat	480
ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttactgcat	540
tctagttgtg gtttgtccaa actcatcaat gtatcttata atgtctgtat accgtcgacc	600
tctagctaga gcttggcgta atcatgggtca tagctgtttc ctgtgtgaaa ttgttatccg	660
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtcctaa	720
tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac	780
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt	840
gggcgtctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga	900
gcggtatcag ctactcaaa ggcgtaata cggttatcca cagaatcagg ggataacgca	960
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg	1020
ctggcgtttt tccataggct ccgccccct gacgagcatc aaaaaaatcg acgtcaagt	1080
cagaggtggc gaaacccgac aggactataa agataaccagg cgtttcccc tggaaagctcc	1140
ctcgtgcgct ctctgtttc gaccctgccg cttaccggat acctgtccgc tttctccct	1200
tcgggaagcg tggcgcttcc tcaatgctca cgctgtaggt atctcagttc ggtgtaggtc	1260
gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta	1320
tccgtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca	1380
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag	1440
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag	1500
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaac caccgctggt	1560
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaagg atctcaagaa	1620
gatcctttga tcttttctac ggggtctgac gctcagtggg acgaaaactc acgttaaggg	1680
atthttgtca tgagattatc aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga	1740
agthtttaa caatctaaag tatatatgag taaacttggc ctgacagtta ccaatgctta	1800
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgctgactc	1860
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg	1920
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga	1980
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt	2040
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt	2100

gctacaggca	tcgtggtgtc	acgctcgtcg	tttgggtatgg	cttcattcag	ctccggttcc	2160
caacgatcaa	ggcgagttac	atgatcccc	atgttgtgca	aaaaagcgg	tagctccttc	2220
ggtcctccga	tcgttgtcag	aagtaagttg	gccgcagtgt	tatcactcat	ggttatggca	2280
gcactgcata	attctcttac	tgtcatgcca	tccgtaagat	gcttttctgt	gactgggtgag	2340
tactcaacca	agtcattctg	agaatagtgt	atgcggcgac	cgagttgctc	ttgcccggcg	2400
tcaatacggg	ataataccgc	gccacatagc	agaacttta	aagtgtcat	cattggaaaa	2460
cgttcttcgg	ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	2520
cccactcgtg	cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	2580
gcaaaaacag	gaaggcaaaa	tgccgcaaaa	aagggaataa	ggcgacacg	gaaatgttga	2640
atactcatac	tcttctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	2700
agcggataca	tatttgaatg	tatttagaaa	aataacaaa	taggggttcc	gcgcacattt	2760
ccccgaaaag	tgccacctga	cgtcgacgga	tcgggagatc	cgatgttttc	tgagttactt	2820
ttgtatcccc	acccccctc	gaggatcgat	ccgctcggag	gacagtactc	cgctcggagg	2880
acagtactcc	gctcggagga	cagtactccg	ctcgaggaca	gtactccgct	cggaggacag	2940
tactccgatc	cgtcgactct	agtcgcctcg	gaggacagta	ctccgctcgg	aggacagtac	3000
tccgctcggg	ggacagtact	ccgctcggag	acagtactcg	attcgaggac	attttgacac	3060
ccccataata	tttttccaga	attaacagta	taaattgcat	ctcttgttca	agagttccct	3120
atcactctct	ttaatcacta	ctcacagtaa	cctcaactcc	tgcccaagct	tgggtaccacc	3180
atggaccag	aaacgctggt	gaaagtaaaa	gatgctgaag	atcagttggg	tgcacgagtg	3240
ggttacatcg	aactggatct	caacagcgg	aagatccttg	agagttttcg	ccccgaagaa	3300
cgtttttcaa	tgatgagcac	ttttaagtt	ctgctatgtg	gcgcggtatt	atcccgtatt	3360
gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	cttgggttgag	3420
tactcaccag	tcacagaaaa	gcattcttac	gatggcatga	cagtaagaga	attatgcagt	3480
gctgccataa	ccatgagtga	taacactgcg	gccaaacttac	ttctgacaac	gatcggagga	3540
ccgaaggagc	taaccgcttt	tttgacaaac	atgggggatc	atgtaactcg	ccttgatcgt	3600
tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	3660
gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	agcttcccgg	3720
caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	3780
cttccggctg	gctggtttat	tgtgataaaa	tctggagccg	gtgagcgtgg	gtctcgcggt	3840
atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	tcgtagttat	ctacacgacg	3900
gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	tgctcactg	3960
attaagcatt	ggtaatctag	agagggcccg	tttaaaccgg	ctgatcagcc	tcgactgtgc	4020

cttctagttg	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	accctggaag	4080
gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcacgcac	tgtctgagta	4140
ggtgtcattc	tattctgggg	ggtgggggtg	ggcaggacag	caagggggag	gattgggaag	4200
acaatagcag	gcatgctggg	gatgcggtgg	gctctatggc	ttctgaggcg	gaaagaacca	4260
gctggggctc	taggggggat	ccccacgcgc	cctgtagcgg	cgcatthaag	gcggcgggtg	4320
tggtgggttac	gcgcagcgtg	accgctacac	ttgccagcgc	cctagcgccc	gtccttttcg	4380
ctttcttccc	ttcctttctc	gccacgttcg	ccggctttcc	ccgtcaagct	ctaaatcggg	4440
gcatcccttt	agggttccga	tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	4500
aggggtgatg	ttcacgtagt	gggccatcgc	cctgatagac	ggtttttcgc	cctttgacgt	4560
tggagtccac	gttctttaat	agtggactct	tgttccaaac	tggaacaaca	ctcaacccta	4620
tctcgggtcta	ttcttttgat	ttataaggga	ttttggggat	ttcggcctat	tggttaaaaa	4680
atgagctgat	ttaacaaaaa	tttaacgcga	attaattctg	tggaatgtgt	gtcagttagg	4740
gtgtggaaag	tccccaggct	ccccaggcag	gcagaagtat	gcaaagcatg	catctcaatt	4800
agtcagcaac	cagggtgtgga	aagtccccag	gctccccagc	aggcagaagt	atgcaaagca	4860
tgcactctcaa	ttagtcagca	accatagtcc	cgcccctaac	tccgcccata	ccgcccctaa	4920
ctccgcccag	ttccgcccac	tctccgcccc	atggctgact	aatttttttt	atttatgcag	4980
aggccgaggc	cgcctctgcc	tctgagctat	tccagaagta	gtgaggaggc	ttttttggag	5040
gcctaggctt	ttgcaaaaag	ctcccgggag	cttgtatatc	catttttcgga	tctgatcagc	5100
acgtgttgac	aattaatcat	cggcatagta	tatcggcata	gtataatacg	acaagggtgag	5160
gaactaaacc	atggccaagt	tgaccagtgc	cgttcgggtg	ctcaccgcgc	gcgacgtc	5218